

**Search Notes**

Application No.

09/936,697

Examiner

Christopher Nichols, Ph.D.

Applicant(s)

BURNOL ET AL.

Art Unit

1647

**SEARCHED**

Class	Subclass	Date	Examiner

**INTERFERENCE SEARCHED**

Class	Subclass	Date	Examiner

**SEARCH NOTES  
(INCLUDING SEARCH STRATEGY)**

	DATE	EXMR
SEQ ID NO: 5 SEQ ID NO: 6	1/8/2004	CJN

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2004, 10:24:51 ; Search time 40 Seconds  
(without alignments)  
170.631 Million cell updates/sec

Title: US-09-936-697-5  
Perfect score: 43  
Sequence: 1 PMRSISENSLVAMDFSGQKS.....ENPTEALSVAVEEGLAWRKK 43

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 12001

Minimum DB seq length: 42

Maximum DB seq length: 44

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	% Query		Length	DB	ID	Description
	Score	Match				
1	43	100.0	43	21	AAB18941	Peptide derived fr
2	19	44.2	43	21	AAB18937	Peptide derived fr
3	13	30.2	43	21	AAB18945	Peptide derived fr
4	13	30.2	43	21	AAB18949	Peptide derived fr
5	8	18.6	43	21	AAB18953	Peptide derived fr
6	8	18.6	43	21	AAB18957	Peptide derived fr
7	8	18.6	43	21	AAB18961	Peptide derived fr
8	5	11.6	42	21	AAG28134	Arabidopsis thalia
9	5	11.6	42	22	AAO13735	Human polypeptide
10	5	11.6	43	21	AAG26059	Zea mays protein f
11	5	11.6	43	22	AAU20290	Human novel endocr
12	5	11.6	43	24	ABJ18436	Breast specific re
13	5	11.6	44	11	AAR07283	Smooth muscle myos
14	5	11.6	44	22	AAM85983	Human immune/haema
15	4	9.3	42	15	AAR54002	Characteristic pro
16	4	9.3	42	16	AAR66443	PCLUS 6.1-18 (826-
17	4	9.3	42	16	AAR66438	PCLUS 6.1-18 (826-
18	4	9.3	42	17	AAR89344	Cdk2-interacting p
19	4	9.3	42	18	AAW32119	Interaction trap s
20	4	9.3	42	20	AAY05349	HIV-1 CLUVAC pepti
21	4	9.3	42	20	AAY05342	HIV-1 CLUVAC pepti
22	4	9.3	42	21	AAB30449	Amino acid sequenc
23	4	9.3	42	21	AAB53797	Human colon cancer
24	4	9.3	42	21	AAB57234	Human prostate can
25	4	9.3	42	21	AAB38361	Human secreted pro
26	4	9.3	42	21	AAG15587	Arabidopsis thalia
27	4	9.3	42	21	AAG28673	Arabidopsis thalia
28	4	9.3	42	21	AAY76125	Human secreted pro
29	4	9.3	42	21	AAY49338	Cdk2 interacting p
30	4	9.3	42	22	ABG47664	Human liver peptid
31	4	9.3	42	22	ABG49377	Human liver peptid
32	4	9.3	42	22	ABG52150	Human liver peptid
33	4	9.3	42	22	ABG52762	Human liver peptid
34	4	9.3	42	22	ABG01077	Novel human diagno
35	4	9.3	42	22	ABB27642	Human peptide #293
36	4	9.3	42	22	ABB28703	Peptide #1354 enco
37	4	9.3	42	22	ABB29383	Peptide #2034 enco
38	4	9.3	42	22	ABB32081	Peptide #4732 enco
39	4	9.3	42	22	ABB32812	Peptide #318 encod
40	4	9.3	42	22	ABB33885	Peptide #1391 enco
41	4	9.3	42	22	ABB34556	Peptide #2062 enco
42	4	9.3	42	22	ABB37335	Peptide #4841 enco
43	4	9.3	42	22	ABB37914	Peptide #5420 enco
44	4	9.3	42	22	ABB15213	Human nervous syst
45	4	9.3	42	22	ABB15575	Human nervous syst

# ALIGNMENTS

RESULT 1

AAB18941

ID AAB18941 standard; peptide; 43 AA.

XX

AC AAB18941;

XX

DT 08-FEB-2001 (first entry)

XX

DE Peptide derived from the PIR or PIR-SH2 domain of Grb7 protein.

XX

KW Phosphorylated insulin receptor interacting region; Grb7 family protein;  
KW insulin receptor; tyrosine kinase; insulin; insulin-associated disease;  
KW diabetes; obesity; polycystic ovarian syndrome; syndrome X.

XX

OS Homo sapiens.

XX

PN WO200055634-A1.

XX

PD 21-SEP-2000.

XX

PF 14-MAR-2000; 2000WO-FR00613.

XX

PR 15-MAR-1999; 99FR-0003159.

XX

PA (CNRS ) CNRS CENT NAT RECH SCI.

XX

PI Burnol A, Perdereau D, Kasus-Jacobi A, Bereziat V, Girard J;

XX

DR WPI; 2000-587566/55.

XX

PT Fragments of Grb family proteins to identify compounds are useful in  
PT treating insulin-associated diseases, particularly diabetes and obesity

PT -

XX

PS Claim 2; Page 25; 46pp; French.

XX

CC B18937-64 represent the PIR (phosphorylated insulin receptor interacting  
CC region) or PIR-SH2 (Src homology 2) domains of a Grb7 family protein.

CC PIR is the actual binding region but its effect is about 10 times  
CC greater in presence of SH2 (which by itself is inactive). Agents that  
CC affect binding between the peptides and the insulin receptor can  
CC stimulate or inhibit tyrosine kinase activity of the receptor. The  
CC peptides are used for screening molecules for ability to treat diseases  
CC in which insulin is implicated. The peptides are used to identify agents  
CC that are potentially useful for treating insulin-associated diseases,  
CC particularly diabetes and obesity but also polycystic ovarian syndrome  
CC and syndrome X.

XX

SQ Sequence 43 AA;

Query Match 100.0%; Score 43; DB 21; Length 43;

Best Local Similarity 100.0%; Pred. No. 2.2e-39;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PMRSISENSLVAMDFSGQKSRVIENPTEALSVAVEEGLAWRKK 43

|||||

Db 1 PMRSISENSLVAMDFSGQKSRVIENPTEALSVAVEEGLAWRKK 43



RESULT 2

AAB18937

ID AAB18937 standard; peptide; 43 AA.

XX

AC AAB18937;

XX

DT 08-FEB-2001 (first entry)

XX

DE Peptide derived from the PIR or PIR-SH2 domain of Grb7 protein.

XX

KW Phosphorylated insulin receptor interacting region; Grb7 family protein;

KW insulin receptor; tyrosine kinase; insulin; insulin-associated disease;

KW diabetes; obesity; polycystic ovarian syndrome; syndrome X.

XX

OS Rattus sp.

XX

PN WO200055634-A1.

XX

PD 21-SEP-2000.

XX

PF 14-MAR-2000; 2000WO-FR00613.

XX

PR 15-MAR-1999; 99FR-0003159.

XX

PA (CNRS ) CNRS CENT NAT RECH SCI.

XX

PI Burnol A, Perdereau D, Kasus-Jacobi A, Bereziat V, Girard J;

XX

DR WPI; 2000-587566/55.

XX

PT Fragments of Grb family proteins to identify compounds are useful in

PT treating insulin-associated diseases, particularly diabetes and obesity

PT -

XX

PS Claim 2; Page 23; 46pp; French.

XX

CC B18937-64 represent the PIR (phosphorylated insulin receptor interacting  
CC region) or PIR-SH2 (Src homology 2) domains of a Grb7 family protein.

CC PIR is the actual binding region but its effect is about 10 times  
CC greater in presence of SH2 (which by itself is inactive). Agents that

CC affect binding between the peptides and the insulin receptor can

CC stimulate or inhibit tyrosine kinase activity of the receptor. The

CC peptides are used for screening molecules for ability to treat diseases

CC in which insulin is implicated. The peptides are used to identify agents

CC that are potentially useful for treating insulin-associated diseases,

CC particularly diabetes and obesity but also polycystic ovarian syndrome

CC and syndrome X.

XX

SQ Sequence 43 AA;

Query Match 44.2%; Score 19; DB 21; Length 43;

Best Local Similarity 100.0%; Pred. No. 3.1e-13;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 NPTEALSVAVEEGLAWRKK 43

Db

|||||  
25 NPTEALSVAVEEGLAWRKK 43

RESULT 3

AAB18945

ID AAB18945 standard; peptide; 43 AA.

XX

AC AAB18945;

XX

DT 08-FEB-2001 (first entry)

XX

DE Peptide derived from the PIR or PIR-SH2 domain of Grb7 protein.

XX

KW Phosphorylated insulin receptor interacting region; Grb7 family protein;

KW insulin receptor; tyrosine kinase; insulin; insulin-associated disease;

KW diabetes; obesity; polycystic ovarian syndrome; syndrome X.

XX

OS Mus muris.

XX

PN WO200055634-A1.

XX

PD 21-SEP-2000.

XX

PF 14-MAR-2000; 2000WO-FR00613.

XX

PR 15-MAR-1999; 99FR-0003159.

XX

PA (CNRS ) CNRS CENT NAT RECH SCI.

XX

PI Burnol A, Perdereau D, Kasus-Jacobi A, Bereziat V, Girard J;

XX

DR WPI; 2000-587566/55.

XX

PT Fragments of Grb family proteins to identify compounds are useful in

PT treating insulin-associated diseases, particularly diabetes and obesity

PT -

XX

PS Claim 2; Page 27-28; 46pp; French.

XX

CC B18937-64 represent the PIR (phosphorylated insulin receptor interacting  
CC region) or PIR-SH2 (Src homology 2) domains of a Grb7 family protein.

CC PIR is the actual binding region but its effect is about 10 times

CC greater in presence of SH2 (which by itself is inactive). Agents that

CC affect binding between the peptides and the insulin receptor can

CC stimulate or inhibit tyrosine kinase activity of the receptor. The

CC peptides are used for screening molecules for ability to treat diseases

CC in which insulin is implicated. The peptides are used to identify agents

CC that are potentially useful for treating insulin-associated diseases,

CC particularly diabetes and obesity but also polycystic ovarian syndrome

CC and syndrome X.

XX

SQ Sequence 43 AA;

Query Match 30.2%; Score 13; DB 21; Length 43;

Best Local Similarity 100.0%; Pred. No. 1.1e-06;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy            6 SENSLVAMDFSGQ 18  
              |||||||  
Db            6 SENSLVAMDFSGQ 18

RESULT 4

AAB18949

ID    AAB18949 standard; peptide; 43 AA.

XX

AC    AAB18949;

XX

DT    08-FEB-2001    (first entry)

XX

DE    Peptide derived from the PIR or PIR-SH2 domain of Grb7 protein.

XX

KW    Phosphorylated insulin receptor interacting region; Grb7 family protein;  
KW    insulin receptor; tyrosine kinase; insulin; insulin-associated disease;  
KW    diabetes; obesity; polycystic ovarian syndrome; syndrome X.

XX

OS    Homo sapiens.

XX

PN    WO200055634-A1.

XX

PD    21-SEP-2000.

XX

PF    14-MAR-2000; 2000WO-FR00613.

XX

PR    15-MAR-1999;    99FR-0003159.

XX

PA    (CNRS ) CNRS CENT NAT RECH SCI.

XX

PI    Burnol A,   Perdereau D,   Kasus-Jacobi A,   Bereziat V,   Girard J;

XX

DR    WPI; 2000-587566/55.

XX

PT    Fragments of Grb family proteins to identify compounds are useful in  
PT    treating insulin-associated diseases, particularly diabetes and obesity

PT    -

XX

PS    Claim 2; Page 30; 46pp; French.

XX

CC    B18937-64 represent the PIR (phosphorylated insulin receptor interacting  
CC    region) or PIR-SH2 (Src homology 2) domains of a Grb7 family protein.

CC    PIR is the actual binding region but its effect is about 10 times  
CC    greater in presence of SH2 (which by itself is inactive). Agents that  
CC    affect binding between the peptides and the insulin receptor can  
CC    stimulate or inhibit tyrosine kinase activity of the receptor. The  
CC    peptides are used for screening molecules for ability to treat diseases  
CC    in which insulin is implicated. The peptides are used to identify agents  
CC    that are potentially useful for treating insulin-associated diseases,  
CC    particularly diabetes and obesity but also polycystic ovarian syndrome  
CC    and syndrome X.

XX

SQ    Sequence    43 AA;

Query Match                    30.2%;    Score 13;    DB 21;    Length 43;

Best Local Similarity 100.0%; Pred. No. 1.1e-06;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 SENSLVAMDFSGQ 18  
          |||||||  
Db 6 SENSLVAMDFSGQ 18

RESULT 5

AAB18953

ID AAB18953 standard; peptide; 43 AA.

XX

AC AAB18953;

XX

DT 08-FEB-2001 (first entry)

XX

DE Peptide derived from the PIR or PIR-SH2 domain of Grb7 protein.

XX

KW Phosphorylated insulin receptor interacting region; Grb7 family protein;

KW insulin receptor; tyrosine kinase; insulin; insulin-associated disease;

KW diabetes; obesity; polycystic ovarian syndrome; syndrome X.

XX

OS Rattus sp.

XX

PN WO200055634-A1.

XX

PD 21-SEP-2000.

XX

PF 14-MAR-2000; 2000WO-FR00613.

XX

PR 15-MAR-1999; 99FR-0003159.

XX

PA (CNRS ) CNRS CENT NAT RECH SCI.

XX

PI Burnol A, Perdereau D, Kasus-Jacobi A, Bereziat V, Girard J;

XX

DR WPI; 2000-587566/55.

XX

PT Fragments of Grb family proteins to identify compounds are useful in

PT treating insulin-associated diseases, particularly diabetes and obesity

PT -

XX

PS Claim 2; Page 32; 46pp; French.

XX

CC B18937-64 represent the PIR (phosphorylated insulin receptor interacting  
CC region) or PIR-SH2 (Src homology 2) domains of a Grb7 family protein.

CC PIR is the actual binding region but its effect is about 10 times  
CC greater in presence of SH2 (which by itself is inactive). Agents that

CC affect binding between the peptides and the insulin receptor can

CC stimulate or inhibit tyrosine kinase activity of the receptor. The

CC peptides are used for screening molecules for ability to treat diseases

CC in which insulin is implicated. The peptides are used to identify agents

CC that are potentially useful for treating insulin-associated diseases,

CC particularly diabetes and obesity but also polycystic ovarian syndrome

CC and syndrome X.

XX

SQ Sequence 43 AA;

Query Match 18.6%; Score 8; DB 21; Length 43;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 LVAMDFSG 17  
      |||||||  
Db 10 LVAMDFSG 17

RESULT 6

AAB18957

ID AAB18957 standard; peptide; 43 AA.

XX

AC AAB18957;

XX

DT 08-FEB-2001 (first entry)

XX

DE Peptide derived from the PIR or PIR-SH2 domain of Grb7 protein.

XX

KW Phosphorylated insulin receptor interacting region; Grb7 family protein;

KW insulin receptor; tyrosine kinase; insulin; insulin-associated disease;

KW diabetes; obesity; polycystic ovarian syndrome; syndrome X.

XX

OS Homo sapiens.

XX

PN WO200055634-A1.

XX

PD 21-SEP-2000.

XX

PF 14-MAR-2000; 2000WO-FR00613.

XX

PR 15-MAR-1999; 99FR-0003159.

XX

PA (CNRS ) CNRS CENT NAT RECH SCI.

XX

PI Burnol A, Perdereau D, Kasus-Jacobi A, Bereziat V, Girard J;

XX

DR WPI; 2000-587566/55.

XX

PT Fragments of Grb family proteins to identify compounds are useful in

PT treating insulin-associated diseases, particularly diabetes and obesity

PT -

XX

PS Claim 2; Page 34; 46pp; French.

XX

CC B18937-64 represent the PIR (phosphorylated insulin receptor interacting  
CC region) or PIR-SH2 (Src homology 2) domains of a Grb7 family protein.

CC PIR is the actual binding region but its effect is about 10 times  
CC greater in presence of SH2 (which by itself is inactive). Agents that

CC affect binding between the peptides and the insulin receptor can

CC stimulate or inhibit tyrosine kinase activity of the receptor. The

CC peptides are used for screening molecules for ability to treat diseases

CC in which insulin is implicated. The peptides are used to identify agents

CC that are potentially useful for treating insulin-associated diseases,

CC particularly diabetes and obesity but also polycystic ovarian syndrome

CC and syndrome X.

XX

SQ Sequence 43 AA;

Query Match 18.6%; Score 8; DB 21; Length 43;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 LVAMDFSG 17

|||||||

Db 10 LVAMDFSG 17

# RESULT 7

AAB18961

ID AAB18961 standard; peptide; 43 AA.

XX

AC AAB18961;

XX

DT 08-FEB-2001 (first entry)

XX

DE Peptide derived from the PIR or PIR-SH2 domain of Grb7 protein.

XX

KW Phosphorylated insulin receptor interacting region; Grb7 family protein;  
KW insulin receptor; tyrosine kinase; insulin; insulin-associated disease;  
KW diabetes; obesity; polycystic ovarian syndrome; syndrome X.

XX

OS Mus muris.

XX

PN WO200055634-A1.

XX

PD 21-SEP-2000.

XX

PF 14-MAR-2000; 2000WO-FR00613.

XX

PR 15-MAR-1999; 99FR-0003159.

XX

PA (CNRS ) CNRS CENT NAT RECH SCI.

XX

PI Burnol A, Perdereau D, Kasus-Jacobi A, Bereziat V, Girard J;

XX

DR WPI; 2000-587566/55.

XX

PT Fragments of Grb family proteins to identify compounds are useful in  
PT treating insulin-associated diseases, particularly diabetes and obesity

PT -

XX

PS Claim 2; Page 36; 46pp; French.

XX

CC B18937-64 represent the PIR (phosphorylated insulin receptor interacting  
CC region) or PIR-SH2 (Src homology 2) domains of a Grb7 family protein.  
CC PIR is the actual binding region but its effect is about 10 times  
CC greater in presence of SH2 (which by itself is inactive). Agents that  
CC affect binding between the peptides and the insulin receptor can  
CC stimulate or inhibit tyrosine kinase activity of the receptor. The  
CC peptides are used for screening molecules for ability to treat diseases  
CC in which insulin is implicated. The peptides are used to identify agents  
CC that are potentially useful for treating insulin-associated diseases,

CC particularly diabetes and obesity but also polycystic ovarian syndrome  
CC and syndrome X.

XX

SQ Sequence 43 AA;

Query Match 18.6%; Score 8; DB 21; Length 43;

Best Local Similarity 100.0%; Pred. No. 0.3;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 LVAMDFSG 17

|||||||

Db 10 LVAMDFSG 17

#### RESULT 8

AAG28134

ID AAG28134 standard; Protein; 42 AA.

XX

AC AAG28134;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 33238.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR	07-MAY-1999;	99US-0132863.
PR	11-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.



PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.

PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 11.6%; Score 5; DB 21; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 KSRVI 23  
 |||||  
 Db 22 KSRVI 26

RESULT 9  
 AAO13735

ID AAO13735 standard; Protein; 42 AA.  
 XX  
 AC AAO13735;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 27627.  
 XX  
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200164835-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 26-FEB-2001; 2001WO-US04927.  
 XX  
 PR 28-FEB-2000; 2000US-0515126.  
 PR 18-MAY-2000; 2000US-0577409.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 DR WPI; 2001-514838/56.  
 DR N-PSDB; AAI93666.  
 XX  
 PT Isolated nucleic acids and polypeptides, useful for preventing  
 PT diagnosing and treating e.g. leukaemia, inflammation and immune  
 PT disorders -  
 XX  
 PS Claim 20; SEQ ID NO 27627; 1399pp + Sequence Listing; English.  
 XX  
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 42 AA;

Query Match 11.6%; Score 5; DB 22; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy            8 NSLVA 12  
              |||||  
Db            34 NSLVA 38

RESULT 10

AAG26059

ID    AAG26059 standard; Protein; 43 AA.

XX

AC    AAG26059;

XX

DT    17-OCT-2000    (first entry)

XX

DE    Zea mays protein fragment SEQ ID NO: 30371.

XX

KW    Protein identification; signal transduction pathway; metabolic pathway;  
KW    hybridisation assay; genetic mapping; gene expression control; promoter;  
KW    termination sequence; corn.

XX

OS    Zea mays subsp. mays.

XX

PN    EP1033405-A2.

XX

PD    06-SEP-2000.

XX

PF    25-FEB-2000; 2000EP-0301439.

XX

PR    25-FEB-1999;    99US-0121825.

PR    05-MAR-1999;    99US-0123180.

PR    09-MAR-1999;    99US-0123548.

PR    23-MAR-1999;    99US-0125788.

PR    25-MAR-1999;    99US-0126264.

PR    29-MAR-1999;    99US-0126785.

PR    01-APR-1999;    99US-0127462.

PR    06-APR-1999;    99US-0128234.

PR    08-APR-1999;    99US-0128714.

PR    16-APR-1999;    99US-0129845.

PR    19-APR-1999;    99US-0130077.

PR    21-APR-1999;    99US-0130449.

PR    23-APR-1999;    99US-0130510.

PR    23-APR-1999;    99US-0130891.

PR    28-APR-1999;    99US-0131449.

PR    30-APR-1999;    99US-0132048.

PR    30-APR-1999;    99US-0132407.

PR    04-MAY-1999;    99US-0132484.

PR    05-MAY-1999;    99US-0132485.

PR    06-MAY-1999;    99US-0132486.

PR    06-MAY-1999;    99US-0132487.

PR    07-MAY-1999;    99US-0132863.

PR    11-MAY-1999;    99US-0134256.

PR    14-MAY-1999;    99US-0134218.

PR    14-MAY-1999;    99US-0134219.

PR    14-MAY-1999;    99US-0134221.

PR    14-MAY-1999;    99US-0134370.

PR    18-MAY-1999;    99US-0134768.

PR    19-MAY-1999;    99US-0134941.

PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.

PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.

PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 11.6%; Score 5; DB 21; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 NPTEA 29  
 |||||  
 Db 7 NPTEA 11

# RESULT 11

AAU20290

ID AAU20290 standard; Protein; 43 AA.

XX

AC AAU20290;

XX

DT 17-DEC-2001 (first entry)

XX

DE Human novel endocrine antigen, SEQ ID No 347.

XX

KW Human; endocrine antigen; cytostatic; antiinfertility; antidiabetic;  
KW thyroid-active; adrenal-active; androgenic; gastric; gene therapy;  
KW antisense-therapy; antibody; endocrine disorder; hormone imbalance;  
KW reproductive disorder; endocrine cancer; pancreatic disorder;  
KW diabetes mellitus; adrenal gland disorder; hirsutism; thyroid disorder;  
KW hyperthyroidism; hypothalamic disorder; vanishing testes syndrome.

XX

OS Homo sapiens.

XX

PN WO200155319-A2.

XX

PD 02-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US01335.

XX

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226868.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.



PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Barash SC, Ruben SM;

XX

DR WPI; 2001-457726/49.

DR N-PSDB; AAS32393.

XX

PT Isolated polypeptide for treating, preventing and prognosing disorders  
PT related to the endocrine system including endocrine disorders,  
PT reproductive disorders, and gastrointestinal disorders and also for  
PT testing and detection e.g. diagnosis -

XX

PS Claim 11; SEQ ID No 347; 558pp; English.

XX

CC The invention relates to cDNAs encoding novel human endocrine  
CC antigens or a fragment having biological activity, a domain, an epitope,  
CC full length protein, variant, allelic variant or a species homologue of  
CC the cDNA/antigen. The DNAs and polypeptides are useful for preventing,  
CC treating or ameliorating a medical condition when administered

CC (e.g. by gene therapy or antisense-therapy). Identifying mutations in  
CC the genes coding for the antigens is useful for diagnosing a pathological  
CC condition or a susceptibility to a pathological condition. The DNAs,  
CC antigens and antibodies raised against the antigens useful for treating,  
CC preventing and/ or prognosing disorders related to the endocrine system  
CC or hormone imbalance or reproductive disorders, cancers of endocrine  
CC tissues, disorders of the pancreas (e.g. diabetes mellitus), the adrenal  
CC glands (e.g. hirsutism), ovaries, the thyroid (e.g. hyperthyroidism), the  
CC hypothalamus and testes (e.g. vanishing testes syndrome), many examples  
CC of diseases and disorders are given in the specification. The  
CC present sequence represents an endocrine antigen of the invention.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 43 AA;

Query Match 11.6%; Score 5; DB 22; Length 43;  
Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 QKSRV 22

||||

Db 39 QKSRV 43

#### RESULT 12

ABJ18436

ID ABJ18436 standard; Protein; 43 AA.

XX

AC ABJ18436;

XX

DT 30-JAN-2003 (first entry)

XX

DE Breast specific related amino acid sequence SEQ ID No 245.

XX

KW Cytostatic; BSP-agonist; BSP-antagonist; vaccine; gene therapy; cancer;  
KW metastatic; breast cancer; breast specific; human. }

XX

OS Homo sapiens.

XX

PN WO200277232-A2.

XX

PD 03-OCT-2002.

XX

PF 21-NOV-2001; 2001WO-US43815.

XX

PR 22-NOV-2000; 2000US-252509P.

XX

PA (DIAD-) DIADEXUS INC.

XX

PI Salceda S, Macina RA, Recipon H, Pluta J, Sun Y, Liu C;

XX

DR WPI; 2003-018927/01.

XX

PT New isolated nucleic acid molecule, useful for treating breast cancer,

PT and diagnosing or monitoring the presence of metastases of breast  
PT cancer in a patient -  
XX  
PS Claim 11; Page 361; 377pp; English.  
XX  
CC The invention relates to a novel isolated nucleic acid molecule  
CC comprising: a sequence encoding a sequence comprising 11-1518 amino  
CC acids; a sequence comprising 190-8144 bp; or a sequence that selectively  
CC hybridises to, or having at least 60% identity with the 11-1518 amino  
CC acid or 190-8144 nt sequence. The polypeptide and the nucleic acid are  
CC useful for treating breast cancer, and diagnosing or monitoring the  
CC presence of metastases of breast cancer in a patient. The polynucleotides  
CC of the invention can be used to treat disorders by gene therapy. This  
CC sequence represents a breast specific related polypeptide of the  
CC invention.  
XX  
SQ Sequence 43 AA;

Query Match 11.6%; Score 5; DB 24; Length 43;  
Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RSISE 7  
|||||  
Db 26 RSISE 30

# RESULT 13

AAR07283

ID AAR07283 standard; peptide; 44 AA.

XX

AC AAR07283;

XX

DT 31-JAN-1991 (first entry)

XX

DE Smooth muscle myosin-1 immunogen for antibody prodn.

XX

KW Smooth muscle myosin isoform 1; monoclonal antibody; immunogen;

KW diagnosis; arteriosclerosis.

XX

OS Homo sapiens.

XX

PN WO9011520-A.

XX

PD 04-OCT-1990.

XX

PF 26-MAR-1990; 90WO-JP00398.

XX

PR 28-MAR-1989; 89JP-0075884.

XX

PA (YAMS ) YAMASA SHOYU KK.

XX

PI Nagai R, Kuroo M, Kato H;

XX

DR WPI; 1990-320366/42.

XX

PT Antibody against heavy chain of smooth muscle myosin - as reagent

PT for histological staining of smooth muscle or diagnosis of blood  
 PT vessel disorders  
 XX  
 PS Claim 7; Page 42; 61pp; Japanese.  
 XX  
 CC The oligopeptide is used as immunogen for the prodn. of monoclonal  
 CC antibodies recognising isoform SM-1 of the heavy chain of smooth  
 CC muscle myosin, pref. from heart or skeletal muscle, esp. human.  
 CC The peptide contains the part which differs between isoforms SM1-3.  
 CC The antibodies may be obtained by immunisation with the immunogen,  
 CC followed by cell fusion to produce a hybridoma, cloning and  
 CC culturing the chosen hybridoma clone. The Ab is a reagent for  
 CC the histological staining of smooth muscle, and is useful in the  
 CC diagnosis of arteriosclerosis, blood vessel disorders etc.  
 CC See also AAR07284-5.  
 XX  
 SQ Sequence 44 AA;

Query Match 11.6%; Score 5; DB 11; Length 44;  
 Best Local Similarity 100.0%; Pred. No. 5.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 RVIEN 25  
 |||||  
 Db 18 RVIEN 22

#### RESULT 14

AAM85983

ID AAM85983 standard; Protein; 44 AA.

XX

AC AAM85983;

XX

DT 07-NOV-2001 (first entry)

XX

DE Human immune/haematopoietic antigen SEQ ID NO:13576.

XX

KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
 KW cytostatic; gene therapy; vaccine; metastasis.

XX

OS Homo sapiens.

XX

PN WO200157182-A2.

XX

PD 09-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US01354.

XX

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.

PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Barash SC, Ruben SM;

XX

DR WPI; 2001-483426/52.

DR N-PSDB; AAK58764.

XX

PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -

XX

PS Claim 11; SEQ ID NO 13576; 3071pp + Sequence Listing; English.

XX

CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patients own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention.

XX

SQ Sequence 44 AA;

Query Match 11.6%; Score 5; DB 22; Length 44;

Best Local Similarity 100.0%; Pred. No. 5.8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 EALSV 32

|||||

Db 7 EALSV 11

RESULT 15



AAR54002

ID AAR54002 standard; peptide; 42 AA.

XX

AC AAR54002;

XX

DT 25-MAR-2003 (updated)

DT 03-NOV-1994 (first entry)

XX

DE Characteristic protein anchor motif.

XX

KW Conjugate vaccine; Streptococcus infection; group B; polysaccharide;  
KW C protein alpha antigen; neonatal sepsis; meningitis; bca.

XX

OS Streptococcus.

XX

PN W09410317-A2.

XX

PD 11-MAY-1994.

XX

PF 02-NOV-1993; 93WO-US10506.

XX

PR 02-NOV-1992; 92US-0968866.

XX

PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.

PA (GEHO ) GEN HOSPITAL CORP.

XX

PI Ausubel FM, Kasper DL, Madoff LC, Michel JL;

XX

DR WPI; 1994-167472/20.

XX

PT New conjugate vaccine protects against group B Streptococcus

PT infection - comprises gp. B Streptococcus polysaccharide

PT conjugated to C protein alpha antigen deriv., useful against e.g.

PT neonatal sepsis and meningitis

XX

PS Disclosure; Page 80; 103pp; English.

XX

CC The sequence of the group B Streptococcus protein C (bca) was used  
CC in a search to find homologous proteins. A class of Gram-positive  
CC surface proteins with a common membrane anchor motif were found to  
CC be homologous. The amino acid compsn. at the C terminus of the  
CC protein is characteristic of a protein membrane anchor.

CC See also AAR53996-4012.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 42 AA;

Query Match 9.3%; Score 4; DB 15; Length 42;

Best Local Similarity 100.0%; Pred. No. 6.8e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 SVAV 34

|||

Db 21 SVAV 24

Search completed: January 7, 2004, 10:25:47

Job time : 41 secs

OM protein - protein search, using sw model

Run on: January 7, 2004, 10:24:53 ; Search time 21 Seconds  
(without alignments)  
86.637 Million cell updates/sec

Title: US-09-936-697-5  
Perfect score: 43  
Sequence: 1 PMRSISENSLVAMDFSGQKS.....ENPTEALSVAVEEGLAWRKK 43

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2265

Minimum DB seq length: 42  
Maximum DB seq length: 44

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	4	9.3	42	1	US-07-651-710A-25	Sequence 25, Appl
2	4	9.3	42	1	US-08-363-311-19	Sequence 19, Appl
3	4	9.3	42	2	US-08-463-288A-19	Sequence 19, Appl
4	4	9.3	42	2	US-08-470-445A-19	Sequence 19, Appl
5	4	9.3	42	2	US-08-462-679-19	Sequence 19, Appl
6	4	9.3	42	2	US-08-466-210A-19	Sequence 19, Appl
7	4	9.3	42	2	US-08-467-147A-19	Sequence 19, Appl
8	4	9.3	42	2	US-08-455-625-31	Sequence 31, Appl
9	4	9.3	42	2	US-08-455-625-36	Sequence 36, Appl
10	4	9.3	42	2	US-08-469-014-19	Sequence 19, Appl
11	4	9.3	42	3	US-08-504-538A-10	Sequence 10, Appl

12	4	9.3	42	3	US-08-455-685-31	Sequence 31, Appl
13	4	9.3	42	3	US-08-455-685-36	Sequence 36, Appl
14	4	9.3	42	3	US-08-060-988A-31	Sequence 31, Appl
15	4	9.3	42	3	US-08-060-988A-36	Sequence 36, Appl
16	4	9.3	42	4	US-09-346-290-19	Sequence 19, Appl
17	4	9.3	42	4	US-08-630-052-10	Sequence 10, Appl
18	4	9.3	42	5	PCT-US93-10506A-19	Sequence 19, Appl
19	4	9.3	42	5	PCT-US93-10506-19	Sequence 19, Appl
20	4	9.3	42	5	PCT-US94-05142-31	Sequence 31, Appl
21	4	9.3	42	5	PCT-US95-09307-10	Sequence 10, Appl
22	4	9.3	43	1	US-08-306-872-1	Sequence 1, Appli
23	4	9.3	43	1	US-07-998-003A-6	Sequence 6, Appli
24	4	9.3	43	1	US-08-453-274B-6	Sequence 6, Appli
25	4	9.3	43	1	US-08-453-695A-6	Sequence 6, Appli
26	4	9.3	43	1	US-08-268-161A-6	Sequence 6, Appli
27	4	9.3	43	2	US-08-453-702A-6	Sequence 6, Appli
28	4	9.3	43	3	US-09-099-639-6	Sequence 6, Appli
29	4	9.3	43	4	US-09-205-258-891	Sequence 891, App
30	4	9.3	43	5	PCT-US93-12588-6	Sequence 6, Appli
31	4	9.3	43	5	PCT-US95-08071-6	Sequence 6, Appli
32	4	9.3	43	5	PCT-US95-10989-1	Sequence 1, Appli
33	4	9.3	44	1	US-08-145-708A-1	Sequence 1, Appli
34	4	9.3	44	1	US-07-956-700B-83	Sequence 83, Appl
35	4	9.3	44	1	US-08-476-537-83	Sequence 83, Appl
36	4	9.3	44	1	US-08-485-607-83	Sequence 83, Appl
37	4	9.3	44	2	US-08-248-839C-63	Sequence 63, Appl
38	4	9.3	44	2	US-08-331-454-1	Sequence 1, Appli
39	4	9.3	44	2	US-08-475-879-83	Sequence 83, Appl
40	4	9.3	44	4	US-09-083-541-6	Sequence 6, Appli
41	4	9.3	44	4	US-09-433-043B-83	Sequence 83, Appl
42	3	7.0	42	1	US-07-641-971B-4	Sequence 4, Appli
43	3	7.0	42	1	US-07-781-248A-4	Sequence 4, Appli
44	3	7.0	42	1	US-07-842-089E-3	Sequence 3, Appli
45	3	7.0	42	1	US-07-651-710A-23	Sequence 23, Appl

# ALIGNMENTS

## RESULT 1

US-07-651-710A-25

; Sequence 25, Application US/07651710A

; Patent No. 5362864

; GENERAL INFORMATION:

; APPLICANT: Chua, Nam-Hai

; TITLE OF INVENTION: Trans-Activating Factor-1

; NUMBER OF SEQUENCES: 45

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

```

;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/07/651,710A
;   FILING DATE:  19910206
;   CLASSIFICATION:  800
;   ATTORNEY/AGENT INFORMATION:
;   NAME:  Misrock, S. Leslie
;   REGISTRATION NUMBER:  30,742
;   REFERENCE/DOCKET NUMBER:  3288-014
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  212 790-9090
;   TELEFAX:  212 8698864/9741
;   TELEX:  66141 PENNIE
;   INFORMATION FOR SEQ ID NO:  25:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:  42 amino acids
;   TYPE:  AMINO ACID
;   STRANDEDNESS:  single
;   TOPOLOGY:  unknown
;   MOLECULE TYPE:  peptide
US-07-651-710A-25

```

```

Query Match          9.3%;  Score 4;  DB 1;  Length 42;
Best Local Similarity 100.0%;  Pred. No. 1.4e+03;
Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

```

```

Qy          6 SENS 9
           ||||
Db          16 SENS 19

```

# RESULT 2

US-08-363-311-19

```

; Sequence 19, Application US/08363311
; Patent No. 6548241
;   GENERAL INFORMATION:
;   APPLICANT:  Michel, James L.
;   APPLICANT:  Ausubel, Frederick M.
;   TITLE OF INVENTION:  Conjugate Vaccine Against Group B
;   TITLE OF INVENTION:  Streptococcus
;   NUMBER OF SEQUENCES:  29
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
;   STREET:  1225 Connecticut Avenue, N.W.
;   CITY:  Washington
;   STATE:  D.C.
;   COUNTRY:  USA
;   ZIP:  20036-2678
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Floppy disk
;   COMPUTER:  IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/363,311
;   FILING DATE:

```

; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/968,866  
; FILING DATE: 02-NOV-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cimbala, Michele A.  
; REGISTRATION NUMBER: 33,851  
; REFERENCE/DOCKET NUMBER: 0609.3740004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 466-0800  
; TELEFAX: (202) 833-8716  
; TELEX: 248636 SSK  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 42 amino acids  
; TYPE: amino acid  
; TOPOLOGY: both  
; MOLECULE TYPE: peptide  
US-08-363-311-19

Query Match 9.3%; Score 4; DB 1; Length 42;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 SVAV 34  
|||  
Db 21 SVAV 24

RESULT 3

US-08-463-288A-19

; Sequence 19, Application US/08463288A  
; Patent No. 5820860  
; GENERAL INFORMATION:  
; APPLICANT: Michel, James L.  
; APPLICANT: Kasper, Dennis L.  
; APPLICANT: Ausubel, Frederick M.  
; APPLICANT: Madoff, Lawrence C.  
; TITLE OF INVENTION: Conjugate Vaccine For Group B  
; TITLE OF INVENTION: Streptococcus  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, NW, Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,288A  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 424

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugaisky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.2370007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
US-08-463-288A-19

```

```

Query Match          9.3%; Score 4; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      31 SVAV 34
        ||||
Db      21 SVAV 24

```

#### RESULT 4

US-08-470-445A-19

```

; Sequence 19, Application US/08470445A
; Patent No. 5843444
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine Against Group B
; TITLE OF INVENTION: Streptococcus
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,445A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugaisky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.237000A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
US-08-470-445A-19

```

```

Query Match          9.3%; Score 4; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      31 SVAV 34
      ||||
Db      21 SVAV 24

```

RESULT 5

US-08-462-679-19

; Sequence 19, Application US/08462679

; Patent No. 5847081

; GENERAL INFORMATION:

; APPLICANT: Michel, James L.

; APPLICANT: Kasper, Dennis L.

; APPLICANT: Ausubel, Frederick M.

; APPLICANT: Madoff, Lawrence C.

; TITLE OF INVENTION: Conjugate Vaccine For Group B

; TITLE OF INVENTION: Streptococcus

; NUMBER OF SEQUENCES: 65

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

; STREET: 1100 New York Avenue, NW, Suite 600

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA



```

; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,679
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugaisky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.2370008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
US-08-462-679-19

```

```

Query Match          9.3%; Score 4; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches    4; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

```

```

Qy      31 SVAV 34
        ||||
Db      21 SVAV 24

```

# RESULT 6

US-08-466-210A-19

```

; Sequence 19, Application US/08466210A
; Patent No. 5858362
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine For Group B
; TITLE OF INVENTION: Streptococcus
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
 ; STREET: 1100 New York Avenue, NW, Suite 600  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20005-3934  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/466,210A  
 ; FILING DATE: 06-JUN-1995  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/363,311  
 ; FILING DATE: 22-DEC-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/968,866  
 ; FILING DATE: 02-NOV-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/408,036  
 ; FILING DATE: 15-SEP-1989  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Bugaisky, Lawrence B.  
 ; REGISTRATION NUMBER: 35,086  
 ; REFERENCE/DOCKET NUMBER: 0609.237000B  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 371-2600  
 ; TELEFAX: (202) 371-2540  
 ; TELEX: 248636 SSK  
 ; INFORMATION FOR SEQ ID NO: 19:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 42 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: both  
 ; MOLECULE TYPE: peptide  
 US-08-466-210A-19

Query Match 9.3%; Score 4; DB 2; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 SVAV 34  
 ||||  
 Db 21 SVAV 24

# RESULT 7

US-08-467-147A-19  
 ; Sequence 19, Application US/08467147A  
 ; Patent No. 5908629  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Michel, James L.  
 ; APPLICANT: Kasper, Dennis L.  
 ; APPLICANT: Ausubel, Frederick M.

```

; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine For Group B
; TITLE OF INVENTION: Streptococcus
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,147A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugaisky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.2370009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
US-08-467-147A-19

```

```

Query Match          9.3%; Score 4; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches    4; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

```

```

Qy      31 SVAV 34
        ||||
Db      21 SVAV 24

```

```

RESULT 8
US-08-455-625-31
; Sequence 31, Application US/08455625

```

```

; Patent No. 5932218
; GENERAL INFORMATION:
;   APPLICANT: Berzofsky, Jay A.
;   APPLICANT: Ahlers, Jeffrey D.
;   APPLICANT: Pendleton, C. D.
;   APPLICANT: Nara, Peter
;   APPLICANT: Shirai, Mutsunori
;   TITLE OF INVENTION: COMPOSITE SYNTHETIC PEPTIDE CONSTRUCT
;   TITLE OF INVENTION: ELICITING NEUTRALIZING ANTIBODIES AND CYTOTOXIC T
;   TITLE OF INVENTION: LYMPHOCYTES AGAINST HIV
;   NUMBER OF SEQUENCES: 36
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Birch, Stewart, Kolasch & Birch
;     STREET: P.O. Box 747
;     CITY: Falls Church
;     STATE: Virginia
;     COUNTRY: USA
;     ZIP: 22040-0747
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/455,625
;     FILING DATE:
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/060,988
;     FILING DATE: 14-MAY-1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Svensson, Leonard R.
;     REGISTRATION NUMBER: 30330
;     REFERENCE/DOCKET NUMBER: 1173-434P
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 703-205-8000
;     TELEFAX: 703-205-8050
;   INFORMATION FOR SEQ ID NO: 31:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 42 amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;     MOLECULE TYPE: peptide
;     FRAGMENT TYPE: internal
;     FEATURE:
;       NAME/KEY: Peptide
;       LOCATION: 1..42
;       OTHER INFORMATION: /label= peptide
;       OTHER INFORMATION: /note= "pclus6.1-18MN peptide, see Table VI"
US-08-455-625-31

```

```

Query Match          9.3%; Score 4; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      21 RVIE 24
      ||||

```

## RESULT 9

US-08-455-625-36

; Sequence 36, Application US/08455625

; Patent No. 5932218

## ; GENERAL INFORMATION:

; APPLICANT: Berzofsky, Jay A.

; APPLICANT: Ahlers, Jeffrey D.

; APPLICANT: Pendleton, C. D.

; APPLICANT: Nara, Peter

; APPLICANT: Shirai, Mutsunori

; TITLE OF INVENTION: COMPOSITE SYNTHETIC PEPTIDE CONSTRUCT

; TITLE OF INVENTION: ELICITING NEUTRALIZING ANTIBODIES AND CYTOTOXIC T

; TITLE OF INVENTION: LYMPHOCYTES AGAINST HIV

; NUMBER OF SEQUENCES: 36

## ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Birch, Stewart, Kolasch &amp; Birch

; STREET: P.O. Box 747

; CITY: Falls Church

; STATE: Virginia

; COUNTRY: USA

; ZIP: 22040-0747

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

## ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/455,625

; FILING DATE:

; CLASSIFICATION: 435

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/060,988

; FILING DATE: 14-MAY-1993

## ; ATTORNEY/AGENT INFORMATION:

; NAME: Svensson, Leonard R.

; REGISTRATION NUMBER: 30330

; REFERENCE/DOCKET NUMBER: 1173-434P

## ; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-205-8000

; TELEFAX: 703-205-8050

## ; INFORMATION FOR SEQ ID NO: 36:

## ; SEQUENCE CHARACTERISTICS:

; LENGTH: 42 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: internal

## ; FEATURE:

; NAME/KEY: Peptide

; LOCATION: 1..42

; OTHER INFORMATION: /label= peptide

; OTHER INFORMATION: /note=pplus6.1-18IIIB peptid

US-08-455-625-36

Query Match 9.3%; Score 4; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 RVIE 24  
|||  
Db 2 RVIE 5

RESULT 10

US-08-469-014-19

; Sequence 19, Application US/08469014

; Patent No. 5968521

; GENERAL INFORMATION:

; APPLICANT: Michel, James L.

; APPLICANT: Kasper, Dennis L.

; APPLICANT: Ausubel, Frederick M.

; APPLICANT: Madoff, Lawrence C.

; TITLE OF INVENTION: Conjugate Vaccine Against Group B

; TITLE OF INVENTION: Streptococcus

; NUMBER OF SEQUENCES: 65

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

; STREET: 1100 New York Avenue, NW, Suite 600

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20005-3934

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/469,014

; FILING DATE: 05-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/363,311

; FILING DATE: 22-DEC-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/968,866

; FILING DATE: 02-NOV-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/408,036

; FILING DATE: 15-SEP-1989

; ATTORNEY/AGENT INFORMATION:

; NAME: Bugaisky, Lawrence B.

; REGISTRATION NUMBER: 35,086

; REFERENCE/DOCKET NUMBER: 0609.2370006

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 371-2600

; TELEFAX: (202) 371-2540

; TELEX: 248636 SSK

; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 42 amino acids

; TYPE: amino acid  
; TOPOLOGY: both  
; MOLECULE TYPE: peptide  
US-08-469-014-19

Query Match 9.3%; Score 4; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 SVAV 34  
|||  
Db 21 SVAV 24

RESULT 11

US-08-504-538A-10

; Sequence 10, Application US/08504538A  
; Patent No. 6004746  
; GENERAL INFORMATION:  
; APPLICANT: Brent, Roger  
; APPLICANT: McCoy, John M.  
; APPLICANT: Jessen, Timm H.  
; TITLE OF INVENTION: INTERACTION TRAP SYSTEMS FOR DETECTING  
; TITLE OF INVENTION: PROTEIN INTERACTIONS  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Clark & Elbing LLP  
; STREET: 176 Federal Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02110-2214  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/504,538A  
; FILING DATE: 07/20/95  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/278,082  
; FILING DATE: 07/20/94  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Paul T. Clark  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 00786/259001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 428-0200  
; TELEFAX: (617) 428-7045  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 42  
; TYPE: amino acid  
; STRANDEDNESS:

; TOPOLOGY: linear  
US-08-504-538A-10

Query Match 9.3%; Score 4; DB 3; Length 42;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 GLAW 40  
|||  
Db 17 GLAW 20

RESULT 12

US-08-455-685-31

; Sequence 31, Application US/08455685

; Patent No. 6214347

; GENERAL INFORMATION:

; APPLICANT: Berzofsky, Jay A.

; APPLICANT: Ahlers, Jeffrey D.

; APPLICANT: Pendleton, C. David

; APPLICANT: Nara, Peter

; APPLICANT: Shirai, Mutsunori

; TITLE OF INVENTION: MULTIDETERMINANT PEPTIDES THAT ELICIT

; TITLE OF INVENTION: HELPER T-LYMPHOCYTE, CYTOTOXIC T LYMPHOCYTE AND

; TITLE OF INVENTION: NEUTRALIZING ANTIBODY RESPONSES AGAINST HIV-1

; NUMBER OF SEQUENCES: 40

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows95

; SOFTWARE: FastSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/455,685

; FILING DATE: 31-MAY-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/060,988

; FILING DATE: 14-MAY-1993

; APPLICATION NUMBER: 07/847,311

; FILING DATE: 06-MAR-1992

; APPLICATION NUMBER: 07/751,998

; FILING DATE: 29-AUG-1991

; APPLICATION NUMBER: 07/148,692

; FILING DATE: 26-JAN-1988

; ATTORNEY/AGENT INFORMATION:

; NAME: Beattie, Ingrid A.

; REGISTRATION NUMBER: P-42,306

; REFERENCE/DOCKET NUMBER: 08830/022003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906



; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 42 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-455-685-31

Query Match 9.3%; Score 4; DB 3; Length 42;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 RVIE 24  
|||  
Db 2 RVIE 5

RESULT 13

US-08-455-685-36

; Sequence 36, Application US/08455685  
; Patent No. 6214347

; GENERAL INFORMATION:

; APPLICANT: Berzofsky, Jay A.  
; APPLICANT: Ahlers, Jeffrey D.  
; APPLICANT: Pendleton, C. David  
; APPLICANT: Nara, Peter  
; APPLICANT: Shirai, Mutsunori

; TITLE OF INVENTION: MULTIDETERMINANT PEPTIDES THAT ELICIT  
; TITLE OF INVENTION: HELPER T-LYMPHOCYTE, CYTOTOXIC T LYMPHOCYTE AND  
; TITLE OF INVENTION: NEUTRALIZING ANTIBODY RESPONSES AGAINST HIV-1  
; NUMBER OF SEQUENCES: 40

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/455,685  
; FILING DATE: 31-MAY-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/060,988  
; FILING DATE: 14-MAY-1993  
; APPLICATION NUMBER: 07/847,311  
; FILING DATE: 06-MAR-1992  
; APPLICATION NUMBER: 07/751,998  
; FILING DATE: 29-AUG-1991  
; APPLICATION NUMBER: 07/148,692  
; FILING DATE: 26-JAN-1988

; ATTORNEY/AGENT INFORMATION:

; NAME: Beattie, Ingrid A.  
; REGISTRATION NUMBER: P-42,306  
; REFERENCE/DOCKET NUMBER: 08830/022003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 42 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-455-685-36

Query Match 9.3%; Score 4; DB 3; Length 42;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 RVIE 24  
|||  
Db 2 RVIE 5

RESULT 14

US-08-060-988A-31  
; Sequence 31, Application US/08060988A  
; Patent No. 6294322  
; GENERAL INFORMATION:  
; APPLICANT: Berzofsky, Jay A.  
; APPLICANT: Ahlers, Jeffrey D.  
; APPLICANT: Pendleton, C. David  
; APPLICANT: Nara, Peter  
; APPLICANT: Shirai, Mutsunori  
; TITLE OF INVENTION: MULTIDETERMINANT PEPTIDES  
; TITLE OF INVENTION: THAT ELICIT  
; TITLE OF INVENTION: HELPER T-LYMPHOCYTE, CYTOTOXIC T LYMPHOCYTE AND  
; TITLE OF INVENTION: NEUTRALIZING ANTIBODY RESPONSES AGAINST HIV-1  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/060,988A  
; FILING DATE: 14-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/847,311  
; FILING DATE: 06-MAR-1992

```

; APPLICATION NUMBER: 07/751,998
; FILING DATE: 29-AUG-1991
; APPLICATION NUMBER: 07/148,692
; FILING DATE: 26-JAN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Beattie, Ingrid A.
; REGISTRATION NUMBER: P-42,306
; REFERENCE/DOCKET NUMBER: 08830/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-060-988A-31

```

```

Query Match          9.3%; Score 4; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      21 RVIE 24
        ||||
Db      2 RVIE 5

```

# RESULT 15

US-08-060-988A-36

```

; Sequence 36, Application US/08060988A
; Patent No. 6294322
; GENERAL INFORMATION:
; APPLICANT: Berzofsky, Jay A.
; APPLICANT: Ahlers, Jeffrey D.
; APPLICANT: Pendleton, C. David
; APPLICANT: Nara, Peter
; APPLICANT: Shirai, Mutsunori
; TITLE OF INVENTION: MULTIDETERMINANT PEPTIDES
; TITLE OF INVENTION: THAT ELICIT
; TITLE OF INVENTION: HELPER T-LYMPHOCYTE, CYTOTOXIC T LYMPHOCYTE AND
; TITLE OF INVENTION: NEUTRALIZING ANTIBODY RESPONSES AGAINST HIV-1
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/060,988A  
; FILING DATE: 14-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/847,311  
; FILING DATE: 06-MAR-1992  
; APPLICATION NUMBER: 07/751,998  
; FILING DATE: 29-AUG-1991  
; APPLICATION NUMBER: 07/148,692  
; FILING DATE: 26-JAN-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Beattie, Ingrid A.  
; REGISTRATION NUMBER: P-42,306  
; REFERENCE/DOCKET NUMBER: 08830/022001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 42 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-060-988A-36

Query Match 9.3%; Score 4; DB 3; Length 42;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 RVIE 24  
|||  
Db 2 RVIE 5

Search completed: January 7, 2004, 10:27:19  
Job time : 22 secs

OM protein - protein search, using sw model

Run on: January 7, 2004, 10:24:52 ; Search time 20 Seconds  
(without alignments)  
206.763 Million cell updates/sec

Title: US-09-936-697-5  
Perfect score: 43  
Sequence: 1 PMRSISENSLVAMDFSGQKS.....ENPTEALSVAVEEGLAWRKK 43

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 740

Minimum DB seq length: 42  
Maximum DB seq length: 44

Post-processing: Listing first 45 summaries

Database : PIR\_76:\*  
1: pirl:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	5	11.6	42	2	S41210	serine proteinase
2	4	9.3	42	2	S37685	protein IEF SSP 91
3	4	9.3	42	2	A36533	H+-transporting tw
4	4	9.3	42	2	S58571	photosystem I prot
5	4	9.3	42	2	C42799	photosystem I chai
6	4	9.3	42	2	A42598	urease accessory p
7	4	9.3	42	2	H90761	hypothetical prote
8	4	9.3	42	2	T07557	hypothetical prote
9	4	9.3	42	2	A82802	hypothetical prote
10	4	9.3	43	2	S35580	cysteine proteinas
11	4	9.3	43	2	S66678	serpin - wheat (fr
12	4	9.3	43	2	H84145	hypothetical prote
13	4	9.3	43	2	H89887	hypothetical prote

14	4	9.3	43	2	F98089	degenerate transpo
15	4	9.3	44	1	A1RZJ	photosystem I prot
16	4	9.3	44	2	I37286	olfactory receptor
17	4	9.3	44	2	D85363	hypothetical prote
18	3	7.0	42	1	GIPG	gastric inhibitory
19	3	7.0	42	1	GIBO	gastric inhibitory
20	3	7.0	42	2	A34259	cytochrome P450mt4
21	3	7.0	42	2	S35178	cytochrome P450 (c
22	3	7.0	42	2	S22198	cytochrome-c oxida
23	3	7.0	42	2	C47622	nitrogenase (EC 1.
24	3	7.0	42	2	B48301	glutamate-1-semial
25	3	7.0	42	2	C56406	casein kinase I (E
26	3	7.0	42	2	B39880	streptomycin/spect
27	3	7.0	42	2	I51291	aldolase C - chick
28	3	7.0	42	2	B60195	transforming prote
29	3	7.0	42	2	A37907	thyrotropin beta c
30	3	7.0	42	2	I37543	MHC HLA-DR-beta-1
31	3	7.0	42	2	S49217	ribosomal protein
32	3	7.0	42	2	T07077	heat shock protein
33	3	7.0	42	2	I48852	tropomyosin 5 - mo
34	3	7.0	42	2	S70326	calcium-binding pr
35	3	7.0	42	2	B92087	myelin basic prote
36	3	7.0	42	2	A37896	progesterone recep
37	3	7.0	42	2	S51235	puc2B protein - Rh
38	3	7.0	42	2	S51237	puc3B protein - Rh
39	3	7.0	42	2	T07253	photosystem II pro
40	3	7.0	42	2	S00692	probable photosyst
41	3	7.0	42	2	B42177	KRAB-domain-contai
42	3	7.0	42	2	G56271	lpfA 5'-region hyp
43	3	7.0	42	2	A95019	hypothetical prote
44	3	7.0	42	2	D95088	hypothetical prote
45	3	7.0	42	2	B95184	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

S41210

serine proteinase - *Streptomyces fradiae*

C;Species: *Streptomyces fradiae*

C;Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 19-May-2000

C;Accession: S41210

R;Kitadokoro, K.; Tsuzuki, H.; Nakamura, E.; Sato, T.; Teraoka, H.

Eur. J. Biochem. 220, 55-61, 1994

A;Title: Purification, characterization, primary structure, crystallization and preliminary crystallographic study of a serine proteinase from *Streptomyces fradiae* ATCC 14544.

A;Reference number: S41210; MUID:94164173; PMID:8119298

A;Accession: S41210

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-42 <KIT>

C;Superfamily: streptogrisin A

Query Match 11.6%; Score 5; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 TEALS 31  
|||||  
Db 31 TEALS 35

RESULT 2

S37685

protein IEF SSP 9126 - human (fragments)

C;Species: Homo sapiens (man)

C;Date: 09-Dec-1993 #sequence\_revision 17-Nov-1995 #text\_change 29-Aug-1997

C;Accession: S37685

R;Leffers, H.; Madsen, P.; Rasmussen, H.H.; Honore, B.; Andersen, A.H.; Walbum, E.; Vandekerckhove, J.; Celis, J.E.

J. Mol. Biol. 231, 982-998, 1993

A;Title: Molecular cloning and expression of the transformation sensitive epithelial marker stratifin. A member of a protein family that has been involved in the protein kinase C signalling pathway.

A;Reference number: S34753; MUID:93294871; PMID:8515476

A;Accession: S37685

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-42 <LEF>

A;Note: 8-Ala was also found

C;Superfamily: 14-3-3 protein

Query Match 9.3%; Score 4; DB 2; Length 42;

Best Local Similarity 100.0%; Pred. No. 2.6e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 LSVA 33  
|||  
Db 22 LSVA 25

RESULT 3

A36533

H+-transporting two-sector ATPase (EC 3.6.3.14) alpha chain - Thermus aquaticus (fragment)

C;Species: Thermus aquaticus

C;Date: 12-Apr-1991 #sequence\_revision 12-Apr-1991 #text\_change 03-Jun-2002

C;Accession: A36533

R;Yokoyama, K.; Oshima, T.; Yoshida, M.

J. Biol. Chem. 265, 21946-21950, 1990

A;Title: Thermus thermophilus membrane-associated ATPase. Indication of a eubacterial V-type ATPase.

A;Reference number: A36533; MUID:91072404; PMID:2147690

A;Accession: A36533

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-42 <YOK>

C;Superfamily: vacuolar H+-transporting ATPase 69K chain; H+-transporting ATP synthase alpha chain homology

C;Keywords: ATP biosynthesis; hydrolase

Query Match 9.3%; Score 4; DB 2; Length 42;

Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 EEGL 38  
    ||||  
Db 32 EEGL 35

RESULT 4

S58571

photosystem I protein psaJ - maize chloroplast

C;Species: chloroplast Zea mays (maize)

C;Date: 29-Nov-1995 #sequence\_revision 19-Jan-1996 #text\_change 26-Aug-1999

C;Accession: S58571

R;Maier, R.M.; Neckermann, K.; Igloi, G.L.; Koessel, H.

J. Mol. Biol. 251, 614-628, 1995

A;Title: Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing.

A;Reference number: S58531; MUID:95395841; PMID:7666415

A;Accession: S58571

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-42 <MAI>

A;Cross-references: EMBL:X86563; NID:g902200; PIDN:CAA60305.1; PID:g902241

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995

C;Genetics:

A;Gene: psaJ

A;Genome: chloroplast

C;Superfamily: photosystem I protein psaJ

C;Keywords: chloroplast; membrane-associated complex; photosystem I

Query Match 9.3%; Score 4; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 LSVA 33  
    ||||  
Db 8 LSVA 11

RESULT 5

C42799

photosystem I chain II - Anabaena variabilis (ATCC 29413) (fragment)

C;Species: Anabaena variabilis

C;Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 03-May-1996

C;Accession: C42799

R;Nyhus, K.J.; Ikeuchi, M.; Inoue, Y.; Whitmarsh, J.; Pakrasi, H.B.

J. Biol. Chem. 267, 12489-12495, 1992

A;Title: Purification and characterization of the photosystem I complex from the filamentous cyanobacterium Anabaena variabilis ATCC 29413.

A;Reference number: A42799; MUID:92316925; PMID:1618755

A;Accession: C42799

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-42 <NYH>



C;Superfamily: photosystem I chain II  
C;Keywords: photosynthesis; photosystem I; thylakoid

Query Match 9.3%; Score 4; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 AVEE 36  
|||  
Db 22 AVEE 25

RESULT 6

A42598

urease accessory protein ureG - Escherichia coli (fragment)

C;Species: Escherichia coli

C;Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 19-May-2000

C;Accession: A42598

R;Collins, C.M.; Gutman, D.M.

J. Bacteriol. 174, 883-888, 1992

A;Title: Insertional inactivation of an Escherichia coli urease gene by IS3411.

A;Reference number: A42598; MUID:92121127; PMID:1310093

A;Contents: 1021

A;Accession: A42598

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-42 <COL>

A;Cross-references: GB:S77596; NID:g243731; PIDN:AAB21161.1; PID:g243732

A;Note: sequence extracted from NCBI backbone (NCBIN:77596, NCBIP:77597)

C;Superfamily: hydrogenase expression/formation protein hypB

Query Match 9.3%; Score 4; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 EALS 31  
|||  
Db 19 EALS 22

RESULT 7

H90761

hypothetical protein ECs1064 [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001

C;Accession: H90761

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Ohtsubo, E.; Nakayama, K.; Murata, T.; Tanaka, M.; Tobe, T.; Iida, T.; Takami, H.; Honda, T.; Sasakawa, C.; Ogasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12.

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: H90761

A;Status: preliminary

A;Molecule type: DNA  
A;Residues: 1-42 <HAY>  
A;Cross-references: GB:BA000007; PIDN:BAB34487.1; PID:g13360524; GSPDB:GN00154  
A;Experimental source: strain O157:H7, substrain RIMD 0509952  
C;Genetics:  
A;Gene: ECs1064

Query Match 9.3%; Score 4; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 EGLA 39  
|||  
Db 16 EGLA 19

RESULT 8

T07557

hypothetical protein 42f - Japanese black pine chloroplast

C;Species: chloroplast *Pinus thunbergiana* (Japanese black pine)

C;Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 18-Aug-2000

C;Accession: T07557

R;Wakasugi, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Sugiura, M.  
Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994

A;Title: Loss of all *ndh* genes as determined by sequencing the entire  
chloroplast genome of the black pine *Pinus thunbergii*.

A;Reference number: Z16030; MUID:95024047; PMID:7937893

A;Accession: T07557

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-42 <WAK>

A;Cross-references: EMBL:D17510; NID:g529643; PIDN:BAA04433.1; PID:g1262718

C;Genetics:

A;Genome: chloroplast

C;Keywords: chloroplast

Query Match 9.3%; Score 4; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 LAWR 41  
|||  
Db 21 LAWR 24

RESULT 9

A82802

hypothetical protein XF0471 [imported] - *Xylella fastidiosa* (strain 9a5c)

C;Species: *Xylella fastidiosa*

C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

C;Accession: A82802

R;anonymous, The *Xylella fastidiosa* Consortium of the Organization for  
Nucleotide Sequencing and Analysis, Sao Paulo, Brazil.

Nature 406, 151-157, 2000

A;Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A;Reference number: A82515; MUID:20365717; PMID:10910347

A;Note: for a complete list of authors see reference number A59328 below

A;Accession: A82802  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-42 <SIM>  
 A;Cross-references: GB:AE003897; GB:AE003849; NID:g9105313; PIDN:AAF83281.1;  
 GSPDB:GN00128; XFSC:XF0471  
 A;Experimental source: strain 9a5c  
 R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.;  
 Alvarenga, R.; Alves, L.M.C.; Araya, J.E.; Baia, G.S.; Baptista, C.S.; Barros,  
 M.H.; Bonaccorsi, E.D.; Bordin, S.; Bove, J.M.; Briones, M.R.S.; Bueno, M.R.P.;  
 Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.; Colauto, N.B.;  
 Colombo, C.; Costa, F.F.; Costa, M.C.R.; Costa-Neto, C.M.; Coutinho, L.L.;  
 Cristofani, M.; Dias-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.;  
 Ferreira, A.J.S.  
 submitted to GenBank, June 2000  
 A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco,  
 M.C.; Frohme, M.; Furlan, L.R.; Garnier, M.; Goldman, G.H.; Goldman, M.H.S.;  
 Gomes, S.L.; Gruber, A.; Ho, P.L.; Hoheisel, J.D.; Junqueira, M.L.; Kemper,  
 E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigret, F.; Lambais, M.R.;  
 Leite, L.C.C.; Lemos, E.G.M.; Lemos, M.V.F.; Lopes, S.A.; Lopes, C.R.; Machado,  
 J.A.; Machado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques,  
 M.V.; Martins, E.A.L.  
 A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.;  
 Miyaki, C.Y.; Monteiro-Vitorello, C.B.; Moon, D.H.; Nagai, M.A.; Nascimento,  
 A.L.T.O.; Netto, L.E.S.; Nhani Jr., A.; Nobrega, F.G.; Nunes, L.R.; Oliveira,  
 M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Paris, A.; Peixoto,  
 B.R.; Pereira, G.A.G.; Pereira Jr., H.A.; Pesquero, J.B.; Quaggio, R.B.;  
 Roberto, P.G.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.;  
 Santelli, R.V.; Sawasaki, H.E.  
 A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da  
 Silveira, J.F.; Silvestri, M.L.Z.; Siqueira, W.J.; de Souza, A.A.; de Souza,  
 A.P.; Terenzi, M.F.; Truffi, D.; Tsai, S.M.; Tsuhako, M.H.; Vallada, H.; Van  
 Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Zago, M.A.; Zatz, M.;  
 Meidanis, J.; Setubal, J.C.  
 A;Reference number: A59328  
 A;Contents: annotation  
 C;Genetics:  
 A;Gène: XF0471

Query Match 9.3%; Score 4; DB 2; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 LSVA 33  
 ||||  
 Db 8 LSVA 11

# RESULT 10

S35580

cysteine proteinase (EC 3.4.22.-) IV - mountain papaya (fragment)

C;Species: Carica pubescens (mountain papaya)

C;Date: 09-Jun-1994 #sequence\_revision 01-Nov-1996 #text\_change 07-May-1999

C;Accession: S35580

R;Walreavens, V.; Jaziri, M.; van Beeumen, J.; Schnek, A.G.; Kleinschmidt, T.;  
 Looze, Y.

Biol. Chem. Hoppe-Seyler 374, 501-506, 1993

A;Title: Isolation and preliminary characterization of the cysteine-proteinases from the latex of *Carica candamarcensis* Hook.  
A;Reference number: S35577; MUID:94030669; PMID:8216902  
A;Accession: S35580  
A;Molecule type: protein  
A;Residues: 1-43 <WAL>  
A;Note: the source is designated as *Carica candamarcensis*  
C;Superfamily: papain  
C;Keywords: cysteine proteinase; hydrolase

Query Match 9.3%; Score 4; DB 2; Length 43;  
Best Local Similarity 100.0%; Pred. No. 2.7e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 WRKK 43  
|||  
Db 7 WRKK 10

#### RESULT 11

S66678

serpin - wheat (fragment)

C;Species: *Triticum aestivum* (common wheat)

C;Date: 05-Dec-1998 #sequence\_revision 05-Dec-1998 #text\_change 29-Jan-1999

C;Accession: S66678

R;Rosenkrands, I.; Hejgaard, J.; Rasmussen, S.K.; Bjorn, S.E.

FEBS Lett. 343, 75-80, 1994

A;Title: Serpins from wheat grain.

A;Reference number: S43652; MUID:94215711; PMID:8163022

A;Accession: S66678

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-43 <ROS>

C;Superfamily: antithrombin III

Query Match 9.3%; Score 4; DB 2; Length 43;  
Best Local Similarity 100.0%; Pred. No. 2.7e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 EALS 31  
|||  
Db 7 EALS 10

#### RESULT 12

H84145

hypothetical protein BH3968 [imported] - *Bacillus halodurans* (strain C-125)

C;Species: *Bacillus halodurans*

C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C;Accession: H84145

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiramata, C.; Nakamura, Y.; Ogasawara, N.; Kuhara, S.; Horikoshi, K.

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*.

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: H84145

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-43 <STO>  
A;Cross-references: GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BAB07687.1;  
GSPDB:GN00137  
A;Experimental source: strain C-125  
C;Genetics:  
A;Gene: BH3968

Query Match 9.3%; Score 4; DB 2; Length 43;  
Best Local Similarity 100.0%; Pred. No. 2.7e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 ALSV 32  
|||  
Db 2 ALSV 5

RESULT 13

H89887

hypothetical protein [imported] - Staphylococcus aureus (strain N315)

C;Species: Staphylococcus aureus

C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C;Accession: H89887

R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Aoki, K.; Nagai, Y.; Lian, J.; Ito, T.; Kanamori, M.; Matsumaru, H.; Maruyama, A.; Murakami, H.; Hosoyama, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Hirakawa, H.; Kuhara, S.; Goto, S.; Yabuzaki, J.; Kanehisa, M.; Yamashita, A.; Oshima, K.; Furuya, K.; Yoshino, C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001

A;Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.

A;Reference number: A89758; MUID:21311952; PMID:11418146

A;Accession: H89887

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-43 <KUR>

A;Cross-references: GB:BA000018; PID:g13700964; PIDN:BAB42260.1; GSPDB:GN00149

A;Experimental source: strain N315

C;Genetics:

A;Gene: SAS036

Query Match 9.3%; Score 4; DB 2; Length 43;  
Best Local Similarity 100.0%; Pred. No. 2.7e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 ALSV 32  
|||  
Db 18 ALSV 21

RESULT 14

F98089

degenerate transposase (orf1) [imported] - Streptococcus pneumoniae (strain R6)

C;Species: Streptococcus pneumoniae

C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 22-Oct-2001

C;Accession: F98089

R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; Estrem, S.; Fritz, L.; Fu, D.J.; Fuller, W.; Geringer, C.; Gilmour, R.; Glass, J.S.; Khoja, H.; Kraft, A.; LaGace, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; McHenney, M.; McLeaster, K.; Mundy, C.; Nicas, T.I.; Norris, F.H.; O'Gara, M.; Peery, R.; Robertson, G.T.; Rockey, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.;

Jaskunas, S.R.; Rosteck Jr., P.R.; Skatrud, P.L.; Glass, J.I.

A;Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.

A;Reference number: A97872; MUID:21429245; PMID:11544234

A;Accession: F98089

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-43 <KUR>

A;Cross-references: GB:AE007317; PIDN:AAL00547.1; PID:g15459424; GSPDB:GN00174

C;Genetics:

A;Gene: IS1381-truncation

Query Match 9.3%; Score 4; DB 2; Length 43;

Best Local Similarity 100.0%; Pred. No. 2.7e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 QKSR 21

||||

Db 37 QKSR 40

#### RESULT 15

A1RZJ

photosystem I protein psaJ - rice chloroplast

C;Species: chloroplast *Oryza sativa* (rice)

C;Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 16-Jul-1999

C;Accession: JQ0246; S05126

R;Shimada, H.; Whittier, R.F.; Hiratsuka, J.; Maeda, Y.; Hirai, A.; Sugiura, M. submitted to JIPID, December 1989

A;Reference number: JQ0200

A;Accession: JQ0246

A;Molecule type: DNA

A;Residues: 1-44 <SHI>

A;Experimental source: cv. Nihonbare

R;Hiratsuka, J.; Shimada, H.; Whittier, R.; Ishibashi, T.; Sakamoto, M.; Mori, M.; Kondo, C.; Honji, Y.; Sun, C.R.; Meng, B.Y.; Li, Y.Q.; Kanno, A.; Nishizawa, Y.; Hirai, A.; Shinozaki, K.; Sugiura, M.

Mol. Gen. Genet. 217, 185-194, 1989

A;Title: The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals.

A;Reference number: S05080; MUID:89364698; PMID:2770692

A;Accession: S05126

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-44 <HIR>

A;Cross-references: EMBL:X15901; NID:g11957; PIDN:CAA33968.1; PID:g12007

A;Experimental source: cv. Nihonbare

C;Genetics:

A;Gene: psaJ

A;Map position: CP64622-64756

A;Genome: chloroplast

C;Superfamily: photosystem I protein psaJ

C;Keywords: chloroplast; photosystem I; transmembrane protein

Query Match 9.3%; Score 4; DB 1; Length 44;

Best Local Similarity 100.0%; Pred. No. 2.7e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 LSVA 33

||||

Db 8 LSVA 11

Search completed: January 7, 2004, 10:26:46

Job time : 22 secs

OM protein - protein search, using sw model

Run on: January 7, 2004, 10:26:18 ; Search time 31 Seconds  
(without alignments)  
279.514 Million cell updates/sec

Title: US-09-936-697-5  
Perfect score: 43  
Sequence: 1 PMRSISENSLVAMDFSGQKS.....ENPTEALSVAVEEGLAWRKK 43

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 747907 seqs, 201509753 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4087

Minimum DB seq length: 42

Maximum DB seq length: 44

Post-processing: Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result		%	Query						
No.	Score	Match	Length	DB	ID				Description



1	5	11.6	43	12	US-10-074-024-347	Sequence 347, App
2	4	9.3	42	9	US-09-810-310-5	Sequence 5, Appli
3	4	9.3	42	9	US-09-864-761-33592	Sequence 33592, A
4	4	9.3	42	9	US-09-864-761-34624	Sequence 34624, A
5	4	9.3	42	9	US-09-864-761-35262	Sequence 35262, A
6	4	9.3	42	9	US-09-864-761-37920	Sequence 37920, A
7	4	9.3	42	9	US-09-864-761-38467	Sequence 38467, A
8	4	9.3	42	9	US-09-864-761-46266	Sequence 46266, A
9	4	9.3	42	9	US-09-925-299-1337	Sequence 1337, Ap
10	4	9.3	42	9	US-09-764-860-492	Sequence 492, App
11	4	9.3	42	10	US-09-925-300-1812	Sequence 1812, Ap
12	4	9.3	42	10	US-09-975-143-10	Sequence 10, Appl
13	4	9.3	42	11	US-09-925-299-1337	Sequence 1337, Ap
14	4	9.3	42	11	US-09-764-891-3104	Sequence 3104, Ap
15	4	9.3	42	11	US-09-892-877-118	Sequence 118, App
16	4	9.3	42	12	US-10-029-386-28276	Sequence 28276, A
17	4	9.3	42	12	US-10-029-386-29939	Sequence 29939, A
18	4	9.3	42	12	US-10-212-872-492	Sequence 492, App
19	4	9.3	42	12	US-10-283-940-7	Sequence 7, Appli
20	4	9.3	42	12	US-10-283-963A-9	Sequence 9, Appli
21	4	9.3	42	15	US-10-074-095-492	Sequence 492, App
22	4	9.3	42	15	US-10-205-428-289	Sequence 289, App
23	4	9.3	42	15	US-10-106-698-8236	Sequence 8236, Ap
24	4	9.3	42	15	US-10-162-538-10	Sequence 10, Appl
25	4	9.3	43	9	US-09-798-789-3	Sequence 3, Appli
26	4	9.3	43	9	US-09-810-310-12	Sequence 12, Appl
27	4	9.3	43	9	US-09-864-761-36609	Sequence 36609, A
28	4	9.3	43	9	US-09-864-761-36950	Sequence 36950, A
29	4	9.3	43	9	US-09-864-761-37202	Sequence 37202, A
30	4	9.3	43	9	US-09-864-761-41495	Sequence 41495, A
31	4	9.3	43	9	US-09-864-761-44813	Sequence 44813, A
32	4	9.3	43	9	US-09-864-761-48399	Sequence 48399, A
33	4	9.3	43	10	US-09-981-289-3	Sequence 3, Appli
34	4	9.3	43	10	US-09-764-877-2038	Sequence 2038, Ap
35	4	9.3	43	11	US-09-764-891-3283	Sequence 3283, Ap
36	4	9.3	43	12	US-09-880-573-6	Sequence 6, Appli
37	4	9.3	43	12	US-10-117-982-482	Sequence 482, App
38	4	9.3	43	12	US-10-262-630-25	Sequence 25, Appl
39	4	9.3	43	12	US-09-933-767-891	Sequence 891, App
40	4	9.3	43	12	US-10-321-857-17	Sequence 17, Appl
41	4	9.3	43	12	US-10-321-857-38	Sequence 38, Appl
42	4	9.3	43	12	US-10-321-857-40	Sequence 40, Appl
43	4	9.3	43	12	US-10-321-857-45	Sequence 45, Appl
44	4	9.3	43	12	US-10-321-857-48	Sequence 48, Appl
45	4	9.3	43	12	US-10-321-857-51	Sequence 51, Appl

#### ALIGNMENTS

#### RESULT 1

US-10-074-024-347

; Sequence 347, Application US/10074024

; Publication No. US20030232975A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC001C1  
; CURRENT APPLICATION NUMBER: US/10/074,024  
; CURRENT FILING DATE: 2002-02-14  
; Prior Application removed - See file Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 879  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 347  
; LENGTH: 43  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-074-024-347

Query Match 11.6%; Score 5; DB 12; Length 43;  
Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 QKSRV 22  
    ||||  
Db 39 QKSRV 43

RESULT 2

US-09-810-310-5

; Sequence 5, Application US/09810310  
; Patent No. US20020044948A1  
; GENERAL INFORMATION:  
; APPLICANT: Khleif, Samir N.  
; APPLICANT: Berzofsky, Jay A.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CO-STIMULATION OF  
; TITLE OF INVENTION: IMMUNOLOGICAL RESPONSES TO PEPTIDE ANTIGENS  
; FILE REFERENCE: 15280-415100US  
; CURRENT APPLICATION NUMBER: US/09/810,310  
; CURRENT FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/189,396  
; PRIOR FILING DATE: 2000-03-15  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: HIV PEPTIDE  
; OTHER INFORMATION: ANTIGEN  
US-09-810-310-5

Query Match 9.3%; Score 4; DB 9; Length 42;  
Best Local Similarity 100.0%; Pred. No. 6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 RVIE 24  
    |||  
Db 2 RVIE 5

RESULT 3

US-09-864-761-33592  
; Sequence 33592, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES  
USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 33592  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:

; OTHER INFORMATION: MAP TO AP000150.1  
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.65  
 ; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.2  
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.94  
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9  
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2  
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1  
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.2  
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4  
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3  
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7  
 ; OTHER INFORMATION: EST\_HUMAN HIT: BF695531.1, EVALUE 1.00e-07  
 ; OTHER INFORMATION: EST\_HUMAN HIT: BE275469.1, EVALUE 9.00e-08  
 US-09-864-761-33592

Query Match 9.3%; Score 4; DB 9; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 6e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SISE 7  
 ||||  
 Db 23 SISE 26

#### RESULT 4

US-09-864-761-34624

; Sequence 34624, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES  
 USEFUL FOR

; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

; FILE REFERENCE: Aeomica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34624
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009487.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
; OTHER INFORMATION: SWISSPROT HIT: P41883, EVALUE 5.00e-17
; OTHER INFORMATION: EST_HUMAN HIT: BE708993.1, EVALUE 5.00e-18
US-09-864-761-34624

```

```

Query Match          9.3%; Score 4; DB 9; Length 42;
Best Local Similarity 100.0%; Pred. No. 6e+03;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      28 EALS 31
      ||||
Db      22 EALS 25

```

# RESULT 5

```

US-09-864-761-35262
; Sequence 35262, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G..
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng

```

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES  
USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 35262  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AP000471.1  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.9  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.7

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8  
; OTHER INFORMATION: SWISSPROT HIT: gi2497597, EVALUE 2.10e+00  
; OTHER INFORMATION: EST\_HUMAN HIT: AU142850.1, EVALUE 1.00e-20  
US-09-864-761-35262

Query Match 9.3%; Score 4; DB 9; Length 42;  
Best Local Similarity 100.0%; Pred. No. 6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 EEGL 38  
|||  
Db 8 EEGL 11

#### RESULT 6

US-09-864-761-37920

; Sequence 37920, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES  
USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 37920
;   LENGTH: 42
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AP000009.2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.98
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.5
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
;   OTHER INFORMATION: EST_HUMAN HIT: BF695531.1, EVALUE 1.00e-07
;   OTHER INFORMATION: SWISSPROT HIT: P57054, EVALUE 1.00e-08
US-09-864-761-37920

```

```

Query Match          9.3%; Score 4; DB 9; Length 42;
Best Local Similarity 100.0%; Pred. No. 6e+03;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      4 SISE 7
      ||||
Db      23 SISE 26

```

#### RESULT 7

```

US-09-864-761-38467
; Sequence 38467, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456

```



```

; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38467
;   LENGTH: 42
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC004587.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.6
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.9
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.7
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.3
;   OTHER INFORMATION: SWISSPROT HIT: P46940, EVALUE 8.00e-19
;   OTHER INFORMATION: EST_HUMAN HIT: BF673726.1, EVALUE 1.00e-17
US-09-864-761-38467

```

```

Query Match          9.3%; Score 4; DB 9; Length 42;
Best Local Similarity 100.0%; Pred. No. 6e+03;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

Qy 35 EEGL 38  
|||  
Db 7 EEGL 10

RESULT 8

US-09-864-761-46266

; Sequence 46266, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES  
USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 46266  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC019219.2  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 18  
; OTHER INFORMATION: SWISSPROT HIT: Q08779, EVALUE 3.70e+00  
US-09-864-761-46266

Query Match 9.3%; Score 4; DB 9; Length 42;  
Best Local Similarity 100.0%; Pred. No. 6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 VIEN 25  
|||  
Db 14 VIEN 17

RESULT 9

US-09-925-299-1337  
; Sequence 1337, Application US/09925299  
; Patent No. US20020055627A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA102  
; CURRENT APPLICATION NUMBER: US/09/925,299  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05883  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1556  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1337  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (2)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (7)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (8)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (12)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (23)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE

; LOCATION: (34)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (40)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-299-1337

Query Match 9.3%; Score 4; DB 9; Length 42;  
Best Local Similarity 100.0%; Pred. No. 6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 AMDF 15  
|||  
Db 30 AMDF 33

RESULT 10

US-09-764-860-492

; Sequence 492, Application US/09764860  
; Patent No. US20020094953A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC008  
; CURRENT APPLICATION NUMBER: US/09/764,860  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1198  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 492  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-860-492

Query Match 9.3%; Score 4; DB 9; Length 42;  
Best Local Similarity 100.0%; Pred. No. 6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 RVIE 24  
|||  
Db 18 RVIE 21

RESULT 11

US-09-925-300-1812

; Sequence 1812, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1812  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (12)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (13)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (26)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-300-1812

Query Match 9.3%; Score 4; DB 10; Length 42;  
Best Local Similarity 100.0%; Pred. No. 6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 LSVA 33  
|||  
Db 20 LSVA 23

RESULT 12

US-09-975-143-10

; Sequence 10, Application US/09975143  
; Patent No. US20020155513A1  
; GENERAL INFORMATION:  
; APPLICANT: HSU, Daniel, K.  
; APPLICANT: LIU, Fu-Tong  
; APPLICANT: DOWLING, Christopher, A.  
; TITLE OF INVENTION: GALECTIN EXPRESSION IS INDUCED IN  
; TITLE OF INVENTION: CIRRHOTIC LIVER AND HEPATOCELLULAR CARCINOMA  
; FILE REFERENCE: DANHSU.001C1  
; CURRENT APPLICATION NUMBER: US/09/975,143  
; CURRENT FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: PCT/US00/08561  
; PRIOR FILING DATE: 2000-03-29  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: nematode  
US-09-975-143-10

Query Match 9.3%; Score 4; DB 10; Length 42;  
Best Local Similarity 100.0%; Pred. No. 6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 DFSG 17

Db            ||||  
             23 DFSG 26

RESULT 13

US-09-925-299-1337

; Sequence 1337, Application US/09925299  
; Publication No. US20030040617A9  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA102  
; CURRENT APPLICATION NUMBER: US/09/925,299  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05883  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1556  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1337  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (2)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (7)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (8)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (12)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (23)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (34)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (40)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-299-1337

Query Match                    9.3%; Score 4; DB 11; Length 42;  
Best Local Similarity    100.0%; Pred. No. 6e+03;  
Matches       4; Conservative    0; Mismatches    0; Indels       0; Gaps       0;

Qy            12 AMDF 15  
             ||||  
Db            30 AMDF 33

RESULT 14

US-09-764-891-3104

; Sequence 3104, Application US/09764891  
 ; Publication No. US20030077808A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PC006  
 ; CURRENT APPLICATION NUMBER: US/09/764,891  
 ; CURRENT FILING DATE: 2001-01-17  
 ; Prior application data removed - consult PALM or file wrapper  
 ; NUMBER OF SEQ ID NOS: 10231  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 3104  
 ; LENGTH: 42  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SITE  
 ; LOCATION: (27)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 US-09-764-891-3104

Query Match 9.3%; Score 4; DB 11; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 6e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 SENS 9  
 ||||  
 Db 9 SENS 12

RESULT 15

US-09-892-877-118

; Sequence 118, Application US/09892877  
 ; Publication No. US20030077809A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ruben et. al.  
 ; TITLE OF INVENTION: 97 Human secreted proteins  
 ; FILE REFERENCE: PZ028P1  
 ; CURRENT APPLICATION NUMBER: US/09/892,877  
 ; CURRENT FILING DATE: 2001-06-28  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/437,658  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-11-10  
 ; NUMBER OF SEQ ID NOS: 461  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 118  
 ; LENGTH: 42  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SITE  
 ; LOCATION: (42)  
 ; OTHER INFORMATION: Xaa equals stop translation  
 US-09-892-877-118

Query Match 9.3%; Score 4; DB 11; Length 42;

Best Local Similarity 100.0%; Pred. No. 6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 SLVA 12

||||

Db 12 SLVA 15

Search completed: January 7, 2004, 10:31:47

Job time : 32 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2004, 10:24:53 ; Search time 34 Seconds  
(without alignments)  
326.361 Million cell updates/sec

Title: US-09-936-697-5  
Perfect score: 43  
Sequence: 1 PMRSISENSLVAMDFSGQKS.....ENPTEALSVAVEEGLAWRKK 43

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3084

Minimum DB seq length: 42  
Maximum DB seq length: 44

Post-processing: Listing first 45 summaries

Database : SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query				
No.	Score	Match Length	DB	ID		Description
<hr/>						

1	5	11.6	44	8	Q9T2T3	Q9t2t3 brassica na
2	4	9.3	42	2	Q49062	Q49062 mycoplasma
3	4	9.3	42	2	Q53518	Q53518 escherichia
4	4	9.3	42	4	Q13497	Q13497 homo sapien
5	4	9.3	42	4	Q9HAX9	Q9hax9 homo sapien
6	4	9.3	42	6	O18844	O18844 canis famil
7	4	9.3	42	8	Q32989	Q32989 pinus thunb
8	4	9.3	42	9	Q94M79	Q94m79 lactococcus
9	4	9.3	42	10	Q8S8W3	Q8s8w3 atropa bell
10	4	9.3	42	11	Q9QWQ4	Q9qwq4 rattus norv
11	4	9.3	42	11	Q61813	Q61813 mus musculu
12	4	9.3	42	11	Q9QVT1	Q9qvt1 mus sp. ret
13	4	9.3	42	13	Q8UUI5	Q8uui5 sparus aura
14	4	9.3	42	16	Q9PG31	Q9pg31 xylella fas
15	4	9.3	42	16	Q8XPN6	Q8xpn6 ralstonia s
16	4	9.3	42	16	Q8XD69	Q8xd69 escherichia
17	4	9.3	42	16	Q8P1P8	Q8plp8 streptococc
18	4	9.3	42	16	Q8P0I5	Q8p0i5 streptococc
19	4	9.3	42	17	Q8U2L8	Q8u2l8 pyrococcus
20	4	9.3	43	5	Q8ISB4	Q8isb4 cryptospori
21	4	9.3	43	5	Q8I0R6	Q8i0r6 cryptospori
22	4	9.3	43	8	Q8WE86	Q8we86 cardinalis
23	4	9.3	43	8	Q8WE78	Q8we78 pipilo eryt
24	4	9.3	43	8	Q8LYE7	Q8lye7 anastrepha
25	4	9.3	43	8	Q9G1Q1	Q9glq1 loxodonta a
26	4	9.3	43	10	Q38928	Q38928 arabidopsis
27	4	9.3	43	10	Q41527	Q41527 triticum ae
28	4	9.3	43	10	Q41672	Q41672 vigna mungo
29	4	9.3	43	10	Q41673	Q41673 vigna mungo
30	4	9.3	43	10	Q38927	Q38927 arabidopsis
31	4	9.3	43	10	Q39781	Q39781 gossypium h
32	4	9.3	43	10	Q38762	Q38762 azolla rubr
33	4	9.3	43	10	Q42781	Q42781 glycine max
34	4	9.3	43	10	Q41997	Q41997 arabidopsis
35	4	9.3	43	10	Q41621	Q41621 trifolium s
36	4	9.3	43	10	Q9S8N4	Q9s8n4 triticum ae
37	4	9.3	43	13	Q9DFA5	Q9dfa5 buteo buteo
38	4	9.3	43	16	Q9K5W9	Q9k5w9 bacillus ha
39	4	9.3	43	16	Q99UU4	Q99uu4 staphylococ
40	4	9.3	43	16	Q9K248	Q9k248 chlamydia p
41	4	9.3	43	16	Q8FCL6	Q8fcl6 escherichia
42	4	9.3	43	16	Q8EF80	Q8ef80 shewanella
43	4	9.3	43	16	Q8EEF7	Q8eef7 shewanella
44	4	9.3	43	16	Q8DNG3	Q8dng3 streptococc
45	4	9.3	44	2	Q9ZFC8	Q9zfc8 lactococcus

#### ALIGNMENTS

##### RESULT 1

Q9T2T3

ID Q9T2T3 PRELIMINARY; PRT; 44 AA.  
AC Q9T2T3;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Chaperonin-60 LS2 fragment (Fragments).  
 OS Brassica napus (Rape).  
 OG Mitochondrion.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Brassica.  
 OX NCBI\_TaxID=3708;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=94302168; PubMed=7913238;  
 RA Cloney L.P., Bekkaoui D.R., Feist G.L., Lane W.S., Hemmingsen S.M.;  
 RT "Brassica napus plastid and mitochondrial chaperonin-60 proteins  
 RT contain multiple distinct polypeptides.";  
 RL Plant Physiol. 105:233-241(1994).  
 DR InterPro; IPR001844; Chaprnin\_Cpn60.  
 DR PROSITE; PS00296; CHAPERONINS\_CPN60; 1.  
 FT NON\_TER 1 1  
 FT NON\_CONS 23 24  
 FT NON\_TER 44 44  
 SQ SEQUENCE 44 AA; 4541 MW; 8443D0C9462AF4F3 CRC64;

Query Match 11.6%; Score 5; DB 8; Length 44;  
 Best Local Similarity 100.0%; Pred. No. 8.9e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 AVEEG 37  
 |||||  
 Db 26 AVEEG 30

## RESULT 2

Q49062

ID Q49062 PRELIMINARY; PRT; 42 AA.  
 AC Q49062;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Uracil-phosphoribosyltransferase (Fragment).  
 OS Mycoplasma capricolum.  
 OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;  
 OC Entomoplasmataceae.  
 OX NCBI\_TaxID=2095;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 27343(Kid);  
 RX MEDLINE=96059641; PubMed=7476192;  
 RA Bork P., Ouzounis C., Casari G., Schneider R., Sander C., Dolan M.,  
 RA Gilbert W., Gillevet P.M.;  
 RT "Exploring the Mycoplasma capricolum genome: a minimal cell reveals  
 RT its physiology.";  
 RL Mol. Microbiol. 16:955-967(1995)..  
 DR EMBL; Z33263; CAA83811.1; -.  
 DR HSSP; P70881; 1I5E.  
 DR InterPro; IPR000836; PRTransferase.  
 DR Pfam; PF00156; Pribosyltran; 1.  
 KW Glycosyltransferase; Transferase.  
 FT NON\_TER 1 1

FT NON\_TER 42 42  
SQ SEQUENCE 42 AA; 4433 MW; AE0A81A044A36C69 CRC64;

Query Match 9.3%; Score 4; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 9.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 VAVE 35  
|||  
Db 36 VAVE 39

RESULT 3

Q53518

ID Q53518 PRELIMINARY; PRT; 42 AA.  
AC Q53518;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE UreG protein (Fragment).  
GN UREG.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92121127; PubMed=1310093;  
RA Collins C.M., Gutman D.M.;  
RT "Insertional inactivation of an Escherichia coli urease gene by  
RT IS3411.";  
RL J. Bacteriol. 174:883-888(1992).  
DR EMBL; S77596; AAB21161.1; -.  
DR InterPro; IPR002894; HypB\_UreG.  
DR Pfam; PF01495; HypB\_UreG; 1.  
FT NON\_TER 1 1  
SQ SEQUENCE 42 AA; 4364 MW; FCC09C2DC270E9AC CRC64;

Query Match 9.3%; Score 4; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 9.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 EALS 31  
|||  
Db 19 EALS 22

RESULT 4

Q13497

ID Q13497 PRELIMINARY; PRT; 42 AA.  
AC Q13497;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Mitotic chromosome condensation protein (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Vogt T., Welsh J., Kullmann F., Stolz W., McClelland M.;  
 RL Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; U46075; AAA86843.1; -.  
 DR InterPro; IPR003439; ABC\_transporter.  
 DR InterPro; IPR003405; SMC\_C.  
 DR Pfam; PF02483; SMC\_C; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 42 42  
 SQ SEQUENCE 42 AA; 4524 MW; 63849458CB31C73D CRC64;

Query Match 9.3%; Score 4; DB 4; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 9.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 SLVA 12  
 ||||  
 Db 21 SLVA 24

#### RESULT 5

##### Q9HAX9

ID Q9HAX9 PRELIMINARY; PRT; 42 AA.  
 AC Q9HAX9;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE Ral guanine nucleotide dissociation stimulator (Fragment).  
 GN RALGDS.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Humphrey D., Kwiatkowska J., Henske E.P., Haines J.L., Halley D.,  
 RA van Slegtenhorst M., Kwiatkowski D.J.;  
 RT "Human ralGDS: cDNA cloning, mapping to 9q34, and genomic structure,  
 RT evaluation as TSC1.";  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF295777; AAG10223.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 42 42  
 SQ SEQUENCE 42 AA; 4308 MW; 71AD54D6C550DFE8 CRC64;

Query Match 9.3%; Score 4; DB 4; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 9.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 ALSV 32  
 ||||  
 Db 36 ALSV 39

RESULT 6

O18844

ID O18844 PRELIMINARY; PRT; 42 AA.  
 AC O18844;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE SA (Fragment).  
 GN SA.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Liu P.-C., Shibuya H., Lubahn D.B., Johnson G.S.;  
 RT "A polymorphic (CT)<sub>n</sub> microsatellite in an intron of the canine SA-  
 RT hypertension locus."  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF022384; AAB86695.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 42 42  
 SQ SEQUENCE 42 AA; 4740 MW; 4723A32A35CD3C92 CRC64;

Query Match 9.3%; Score 4; DB 6; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 9.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 WRKK 43  
 ||||  
 Db 10 WRKK 13

RESULT 7

Q32989

ID Q32989 PRELIMINARY; PRT; 42 AA.  
 AC Q32989;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE ORF42f.  
 OS Pinus thunbergii (Green pine) (Japanese black pine).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.  
 OX NCBI\_TaxID=3350;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92212283; PubMed=1557027;  
 RA Tsudzuki J., Nakashima K., Tsudzuki T., Hiratsuka J., Shibata M.,  
 RA Wakasugi T., Sugiura M.;  
 RT "Chloroplast DNA of black pine retains a residual inverted repeat  
 RT lacking rRNA genes: nucleotide sequences of trnQ, trnK, psbA, trnI and  
 RT trnH and the absence of rps16."  
 RL Mol. Gen. Genet. 232:206-214(1992).  
 RN [2]

RP SEQUENCE FROM N.A.  
RX MEDLINE=95094312; PubMed=8001170;  
RA Tsudzuki J., Ito S., Tsudzuki T., Wakasugi T., Sugiura M.;  
RT "A new gene encoding tRNA pro (GGG) is present in the chloroplast  
RT genome of black pine: a compilation of 32 tRNA genes from black pine  
RT chloroplasts.";  
RL Curr. Genet. 26:153-158(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95024047; PubMed=7937893;  
RA Wakasugi T., Tsudzuki J., Ito S., Nakashima K., Tsudzuki T.,  
RA Sugiura M.;  
RT "Loss of all ndh genes as determined by sequencing the entire  
RT chloroplast genome of the black pine *Pinus thunbergii*.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798(1994).  
DR EMBL; D17510; BAA04433.1; -.  
KW Chloroplast.  
SQ SEQUENCE 42 AA; 4480 MW; 0B5B2669DB4AD8A4 CRC64;

Query Match 9.3%; Score 4; DB 8; Length 42;  
Best Local Similarity 100.0%; Pred. No. 9.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 LAWR 41  
|||  
Db 21 LAWR 24

# RESULT 8

Q94M79

ID Q94M79 PRELIMINARY; PRT; 42 AA.  
AC Q94M79;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical protein.  
OS Lactococcus phage BK5-T.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
OX NCBI\_TaxID=31754;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21235794; PubMed=11336549;  
RA Desiere F., Mahanivong C., Hillier A.J., Chandry P.S., Davidson B.E.,  
RA Brussow H.;  
RT "Comparative Genomics of Lactococcal Phages: Insight from the Complete  
RT Genome Sequence of Lactococcus lactis Phage BK5-T.";  
RL Virology 283:240-252(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Mahanivong C., Boyce J.D., Davidson B.E., Hillier A.J.;  
RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Mahanivong C., Boyce J.D., Davidson B.E., Hillier A.J.;  
RT "Analysis of the sequence, cos site and structural proteins of the  
RT Lactococcus lactis temperate bacteriophage BK5-T.";  
RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AF176025; AAK56841.1; -.  
DR EMBL; AJ245616; CAC80204.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 42 AA; 4808 MW; BA3434ED0B857EB8 CRC64;

Query Match 9.3%; Score 4; DB 9; Length 42;  
Best Local Similarity 100.0%; Pred. No. 9.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 NSLV 11  
|||  
Db 3 NSLV 6

RESULT 9  
Q8S8W3

ID Q8S8W3 PRELIMINARY; PRT; 42 AA.  
AC Q8S8W3;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE PSII reaction center subunit IX.  
GN PSAJ.  
OS Atropa belladonna (Belladonna) (Deadly nightshade).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; lamiids; Solanales; Solanaceae; Atropa.  
OX NCBI\_TaxID=33113;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Ab5p(kan);  
RA Schmitz-Linneweber C., Regel R., Gia Du T., Hupfer H., Herrmann R.G.,  
RA Maier R.M.;  
RT "The nucleotide sequence of the plastid chromosome of Atropa  
RT belladonna (deadly nightshade) and its comparison with that of  
RT Nicotiana tabacum with emphasis on sequence elements relevant for  
RT microevolution."  
RL Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AJ316582; CAC88064.1; -.  
DR InterPro; IPR002453; Beta\_tubulin.  
DR InterPro; IPR002615; PSI\_PsaJ.  
DR Pfam; PF01701; PSI\_PsaJ; 1.  
DR ProDom; PD004198; PSI\_PsaJ; 1.  
DR PROSITE; PS00228; TUBULIN\_B\_AUTOREG; 1.  
KW Chloroplast.  
SQ SEQUENCE 42 AA; 4835 MW; DD68F244CEA49E9D CRC64;

Query Match 9.3%; Score 4; DB 10; Length 42;  
Best Local Similarity 100.0%; Pred. No. 9.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 LSVA 33  
|||  
Db 8 LSVA 11



RESULT 10

Q9QWQ4

ID Q9QWQ4 PRELIMINARY; PRT; 42 AA.  
 AC Q9QWQ4;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Nucleoside diphosphate kinase alpha isoform (Fragment).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Wistar;  
 RA Ishikawa N., Taniguchi-Seto H., Munakata Y., Takagi Y., Shimada N.,  
 RA Kimura N.;  
 RT "Multiple transcripts for rat nucleoside diphosphate kinase alpha  
 RT isoform are structually categorized into two groups that exhibit cell-  
 RT specific expression and distinct translation potential.";  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; D89068; BAA13756.1; -.  
 DR HSSP; P22392; 1NUE.  
 DR InterPro; IPR001564; NDK.  
 DR Pfam; PF00334; NDK; 1.  
 DR ProDom; PD001018; NDK; 1.  
 KW Kinase.  
 FT NON\_TER 42 42  
 SQ SEQUENCE 42 AA; 4878 MW; 1E53A4CBFC390622 CRC64;

Query Match 9.3%; Score 4; DB 11; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 9.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 LVAM 13  
 ||||  
 Db 35 LVAM 38

RESULT 11

Q61813

ID Q61813 PRELIMINARY; PRT; 42 AA.  
 AC Q61813;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Transmembrane glycoprotein (Fragment).  
 GN LY-5.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88330145; PubMed=3417340;  
 RA Tung J.-S., Saga Y., Boyse E.A.;  
 RT "Structural features of Ly-5 glycoproteins of the mouse and

RT counterparts in other mammals.";  
 RL Immunogenetics 28:271-277(1988).  
 DR EMBL; M23241; AAA39460.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 42 42  
 SQ SEQUENCE 42 AA; 4454 MW; EE48006FE1D71E12 CRC64;

Query Match 9.3%; Score 4; DB 11; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 9.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SISE 7  
 ||||  
 Db 27 SISE 30

# RESULT 12

## Q9QVT1

ID Q9QVT1 PRELIMINARY; PRT; 42 AA.  
 AC Q9QVT1;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Retinoic acid receptor ALPHAM403, RARALPHAM403 (Fragment).  
 OS Mus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10095;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95198714; PubMed=7891690;  
 RA Matsui T., Sashihara S.;  
 RT "Tissue-specific distribution of a novel C-terminal truncation  
 RT retinoic acid receptor mutant which acts as a negative repressor in a  
 RT promoter-and cell-type-specific manner."  
 RL Mol. Cell. Biol. 15:1961-1967(1995).  
 DR HSSP; P22932; 3LBD.  
 FT NON\_TER 1 1  
 FT NON\_TER 42 42  
 SQ SEQUENCE 42 AA; 4775 MW; F362FDCE9AD57E6E CRC64;

Query Match 9.3%; Score 4; DB 11; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 9.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RSIS 6  
 ||||  
 Db 14 RSIS 17

# RESULT 13

## Q8UUI5

ID Q8UUI5 PRELIMINARY; PRT; 42 AA.  
 AC Q8UUI5;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Caspase-6 (Fragment).  
GN CASP6.  
OS Sparus aurata (Gilthead sea bream).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;  
OC Sparidae; Sparus.  
OX NCBI\_TaxID=8175;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Head kidney;  
RA Pelegrin P., Mulero V., Meseguer J.;  
RT "Identification and characterisation of a fish caspase-6 cDNA.";  
RL Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AJ428523; CAD21592.1; -.  
DR InterPro; IPR002138; ICE\_p10.  
DR Pfam; PF00655; ICE\_p10; 1.  
DR PROSITE; PS50207; CASPASE\_P10; 1.  
FT NON\_TER 1 1  
FT NON\_TER 42 42  
SQ SEQUENCE 42 AA; 4576 MW; 60D898E822DABADF CRC64;

Query Match 9.3%; Score 4; DB 13; Length 42;  
Best Local Similarity 100.0%; Pred. No. 9.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 VAMD 14  
|||  
Db 8 VAMD 11

#### RESULT 14

Q9PG31

ID Q9PG31 PRELIMINARY; PRT; 42 AA.  
AC Q9PG31;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hypothetical protein Xf0471.  
GN XF0471.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
OC Xanthomonadaceae; Xylella.  
OX NCBI\_TaxID=2371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=9a5c;  
RX MEDLINE=20365717; PubMed=10910347;  
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,  
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,  
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,

RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,  
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*.";  
 RL Nature 406:151-159(2000).  
 DR EMBL; AE003897; AAF83281.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 42 AA; 4690 MW; AAF5B1F2FA8E6DC9 CRC64;

Query Match .9.3%; Score 4; DB 16; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 9.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 LSV A 33  
 ||||  
 Db 8 LSV A 11

# RESULT 15

## Q8XPN6

ID Q8XPN6 PRELIMINARY; PRT; 42 AA.  
 AC Q8XPN6;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein RSp1602.  
 GN RSP1602 OR RS02159.  
 OS *Ralstonia solanacearum* (*Pseudomonas solanacearum*).  
 OG Plasmid megaplasmid.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Ralstoniaceae; *Ralstonia*.  
 OX NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GMI1000;  
 RX MEDLINE=21681879; PubMed=11823852;  
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
 RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,  
 RA Chandler M., Choisine N., Claudel-Renard C., Cunnac S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,  
 RA Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,  
 RA Weissenbach J., Boucher C.A.;  
 RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*.";

RL Nature 415:497-502(2002).  
DR EMBL; AL646086; CAD18753.1; -.  
KW Plasmid; Hypothetical protein; Complete proteome.  
SQ SEQUENCE 42 AA; 4492 MW; AE010801082BA52B CRC64;

Query Match 9.3%; Score 4; DB 16; Length 42;  
Best Local Similarity 100.0%; Pred. No. 9.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 LSVA 33  
    ||||  
Db 17 LSVA 20

Search completed: January 7, 2004, 10:28:07  
Job time : 36 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2004, 10:24:51 ; Search time 11 Seconds  
(without alignments)  
183.832 Million cell updates/sec

Title: US-09-936-697-5  
Perfect score: 43  
Sequence: 1 PMRSISENSLVAMDFSGQKS.....ENPTEALSVAVEEGLAWRKK 43

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 286

Minimum DB seq length: 42  
Maximum DB seq length: 44

Post-processing: Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	4	9.3	42	1 PSAD_ANAVA	P31089 anabaena va
2	4	9.3	42	1 PSAJ_MAIZE	P19443 zea mays (m
3	4	9.3	42	1 YDFB_ECOLI	P29009 escherichia
4	4	9.3	43	1 CC3_CARC	P32956 carica cand
5	4	9.3	43	1 CC4_CARC	P32957 carica cand
6	4	9.3	43	1 PSAJ_OENHO	Q9mtk3 oenothera h
7	4	9.3	44	1 PSAJ_ARATH	P56769 arabidopsis
8	4	9.3	44	1 PSAJ_LOTJA	Q9bbr3 lotus japon
9	4	9.3	44	1 PSAJ_ORYSA	P12192 oryza sativ
10	4	9.3	44	1 PSAJ_SPIOL	P17230 spinacia ol
11	4	9.3	44	1 PSAJ_TOBAC	P12193 nicotiana t
12	3	7.0	42	1 BD13_BOVIN	P46171 bos taurus
13	3	7.0	42	1 DLP1_ORNAN	P82172 ornithorhyn
14	3	7.0	42	1 GIP_BOVIN	P09680 bos taurus
15	3	7.0	42	1 GIP_PIG	P01281 sus scrofa
16	3	7.0	42	1 KC1D_BOVIN	P35508 bos taurus
17	3	7.0	42	1 LPAS_STRMU	P10540 streptococc

18	3	7.0	42	1	MTMI_SCHPO	P10839	schizosacch
19	3	7.0	42	1	PSAJ_GUITH	O78456	guillardia
20	3	7.0	42	1	PSAJ_NEPOL	Q9tkz8	nephroselmi
21	3	7.0	42	1	PSBF_GUITH	O78465	guillardia
22	3	7.0	42	1	PSBJ_CHLVU	P56338	chlorella v
23	3	7.0	42	1	PSBJ_EUGGR	P12229	euglena gra
24	3	7.0	42	1	PSBJ_NEPOL	Q9tky4	nephroselmi
25	3	7.0	42	1	RK34_OLILU	P49164	olisthodisc
26	3	7.0	42	1	TYBB_CYPCA	Q9i954	cyprinus ca
27	3	7.0	42	1	Y281_TREPA	O83305	treponema p
28	3	7.0	42	1	YCX2_PAVLU	P31954	pavlova lut
29	3	7.0	43	1	BAGE_HUMAN	Q13072	homo sapien
30	3	7.0	43	1	BDNF_RAJCL	P25430	raja clavat
31	3	7.0	43	1	CC1_CARCN	P32954	carica cand
32	3	7.0	43	1	DEFI_PALPR	P80407	palomena pr
33	3	7.0	43	1	DHBD_ASPOR	P80402	aspergillus
34	3	7.0	43	1	FRIL_SHEEP	P18686	ovis aries
35	3	7.0	43	1	IM9A_RAT	Q9wv97	rattus norv
36	3	7.0	43	1	NLTP_RAPSA	P29420	raphanus sa
37	3	7.0	43	1	NT3_RAJCL	P25434	raja clavat
38	3	7.0	43	1	NT4_VIPLE	P25436	vipera lebe
39	3	7.0	43	1	PMP2_YEAST	P40975	saccharomyc
40	3	7.0	43	1	PSAX_ANASP	P58566	anabaena sp
41	3	7.0	43	1	PSBF_CYACA	Q9tm21	cyanidium c
42	3	7.0	43	1	PSBF_ODOSI	P49474	odontella s
43	3	7.0	43	1	PSBN_ANASP	Q8yyk1	anabaena sp
44	3	7.0	43	1	PSBN_ARATH	P12172	arabidopsis
45	3	7.0	43	1	PSBN_MAGST	Q8hs28	magnolia st

# ALIGNMENTS

## RESULT 1

### PSAD\_ANAVA

ID PSAD\_ANAVA STANDARD; PRT; 42 AA.

AC P31089;

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Photosystem I reaction center subunit II (Photosystem I 16 kDa polypeptide) (PSI-D) (Fragment).

GN PSAD.

OS Anabaena variabilis.

OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.

OX NCBI\_TaxID=1172;

RN [1]

RP SEQUENCE.

RC STRAIN=PCC 7937 / ATCC 29413;

RX MEDLINE=92316925; PubMed=1618755;

RA Nyhus K.J., Ikeuchi M., Inoue Y., Whitmarsh J., Pakrasi H.B.;

RT "Purification and characterization of the photosystem I complex from the filamentous cyanobacterium Anabaena variabilis ATCC 29413.";

RL J. Biol. Chem. 267:12489-12495(1992).

CC -!- FUNCTION: PSAD CAN FORM COMPLEXES WITH FERREDOXIN AND FERREDOXIN-OXIDOREDUCTASE IN PHOTOSYSTEM I (PS I) REACTION CENTRE.

CC -!- SIMILARITY: BELONGS TO THE PSAD FAMILY.

DR PIR; C42799; C42799.  
 DR InterPro; IPR003685; Psad.  
 DR Pfam; PF02531; Psad; 1.  
 KW Photosynthesis; Photosystem I.  
 FT NON\_TER 42 42  
 SQ SEQUENCE 42 AA; 4433 MW; 0BD36D06EEEECEE76 CRC64;

Query Match 9.3%; Score 4; DB 1; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 AVEE 36  
 ||||  
 Db 22 AVEE 25

## RESULT 2

### PSAJ\_MAIZE

ID PSAJ\_MAIZE STANDARD; PRT; 42 AA.  
 AC P19443;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Photosystem I reaction center subunit IX (PSI-J).  
 GN PSAJ.  
 OS Zea mays (Maize), and  
 OS Triticum aestivum (Wheat).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577, 4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Z.mays;  
 RA Haley J., Bogorad L.;  
 RL Submitted (MAY-1989) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Z.mays;  
 RX MEDLINE=95395841; PubMed=7666415;  
 RA Maier R.M., Neckermann K., Igloi G.L., Koessel H.;  
 RT "Complete sequence of the maize chloroplast genome: gene content,  
 RT hotspots of divergence and fine tuning of genetic information by  
 RT transcript editing."  
 RL J. Mol. Biol. 251:614-628(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=T.aestivum; STRAIN=cv. Chinese Spring;  
 RA Ogihara Y., Isono K., Kojima T., Endo A., Hanaoka M., Shiina T.,  
 RA Terachi T., Utsugi S., Murata M., Mori N., Takumi S., Ikeo K.,  
 RA Gojobori T., Murai R., Murai K., Matsuoka Y., Ohnishi Y., Tajiri H.,  
 RA Tsunewaki K.;  
 RT "Chinese spring wheat (Triticum aestivum L.) chloroplast genome:  
 RT complete sequence and contig clones."  
 RL Plant Mol. Biol. Rep. 18:243-253(2000).  
 CC -!- FUNCTION: May help in the organization of the psaE and psaF



CC subunits (By similarity).

CC -!- SIMILARITY: Belongs to the psaJ family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; J04502; AAA84481.1; -. .

DR EMBL; X86563; CAA60305.1; -. .

DR EMBL; AB042240; BAB47053.1; -. .

DR PIR; S58571; S58571.

DR MaizeDB; 69547; -. .

DR HAMAP; MF\_00522; -. . 1.

DR InterPro; IPR002615; PSI\_PsaJ.

DR Pfam; PF01701; PSI\_PsaJ; 1.

DR ProDom; PD004198; PSI\_PsaJ; 1.

KW Chloroplast; Photosystem I; Photosynthesis; Transmembrane.

FT TRANSMEM 7 27 POTENTIAL.

SQ SEQUENCE 42 AA; 4745 MW; DD29FAE6CAB485F6 CRC64;

Query Match 9.3%; Score 4; DB 1; Length 42;

Best Local Similarity 100.0%; Pred. No. 1.2e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 LSVA 33

|||

Db 8 LSVA 11

### RESULT 3

#### YDFB\_ECOLI

ID YDFB\_ECOLI STANDARD; PRT; 42 AA.

AC P29009; P76166;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein ydfB.

GN YDFB OR B1572.

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI\_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=88289404; PubMed=3041373;

RA Cam K., Bejar S., Gil D., Bouche J.-P.;

RT "Identification and sequence of gene dicB: translation of the

RT division inhibitor from an in-phase internal start.";

RL Nucleic Acids Res. 16:6327-6338(1988).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474 (1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=97251357; PubMed=9097039;  
 RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,  
 RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,  
 RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,  
 RA Motomura K., Nakade S., Nakamura Y., Nashimoto H., Nishio Y.,  
 RA Oshima T., Saito N., Sampei G., Seki Y., Sivasundaram S.,  
 RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,  
 RA Yamamoto Y., Horiuchi T.;  
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome  
 RT corresponding to the 28.0-40.1 min region on the linkage map.";  
 RL DNA Res. 3:363-377 (1996).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X07465; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; AE000253; AAC74645.1; ALT\_INIT.  
 DR EMBL; D90799; BAA15277.1; -.  
 DR EMBL; D90800; BAA15299.1; -.  
 DR PIR; G64912; G64912.  
 DR EcoGene; EG11301; ydfB.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 42 AA; 4688 MW; 2FE608C7512F1624 CRC64;

Query Match 9.3%; Score 4; DB 1; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 EGLA 39  
 ||||  
 Db 16 EGLA 19

#### RESULT 4

CC3\_CARCN  
 ID CC3\_CARCN STANDARD; PRT; 43 AA.  
 AC P32956;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cysteine proteinase III (EC 3.4.22.-) (CC-III) (Fragment).  
 OS Carica candamarcensis.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Caricaceae; Carica.  
 OX NCBI\_TaxID=29731;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Latex;  
 RX MEDLINE=94030669; PubMed=8216902;  
 RA Walreavens V., Jaziri M., van Beeumen J., Schnek A.G.,  
 RA Kleinschmidt T., Looze Y.;  
 RT "Isolation and preliminary characterization of the cysteine-  
 RT proteinases from the latex of Carica candamarcensis Hook.";  
 RL Biol. Chem. Hoppe-Seyler 374:501-506(1993).  
 CC -!- PTM: GLYCOSYLATED.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.  
 DR HSSP; P14080; 1YAL.  
 DR MEROPS; C01.020; -.  
 DR InterPro; IPR000668; Peptidase\_C1.  
 DR InterPro; IPR000169; SHprot\_acsite.  
 DR Pfam; PF00112; Peptidase\_C1; 1.  
 DR ProDom; PD000158; Peptidase\_C1; 1.  
 DR PROSITE; PS00639; THIOI\_PROTEASE\_HIS; PARTIAL.  
 DR PROSITE; PS00640; THIOI\_PROTEASE\_ASN; PARTIAL.  
 DR PROSITE; PS00139; THIOI\_PROTEASE\_CYS; 1.  
 KW Hydrolase; Thiol protease; Glycoprotein.  
 FT ACT\_SITE 25 25 BY SIMILARITY.  
 FT NON\_TER 43 43  
 SQ SEQUENCE 43 AA; 4636 MW; F4C5D2881886E291 CRC64;

Query Match 9.3%; Score 4; DB 1; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 WRKK 43  
 ||||  
 Db 7 WRKK 10

#### RESULT 5

##### CC4\_CARCN

ID CC4\_CARCN STANDARD; PRT; 43 AA.  
 AC P32957;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cysteine proteinase IV (EC 3.4.22.-) (CC-IV) (Fragment).  
 OS Carica candamarcensis.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Caricaceae; Carica.  
 OX NCBI\_TaxID=29731;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Latex;  
 RX MEDLINE=94030669; PubMed=8216902;  
 RA Walreavens V., Jaziri M., van Beeumen J., Schnek A.G.,  
 RA Kleinschmidt T., Looze Y.;  
 RT "Isolation and preliminary characterization of the cysteine-

RT proteinases from the latex of *Carica candamarcensis* Hook.";  
 RL Biol. Chem. Hoppe-Seyler 374:501-506(1993).  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.  
 DR HSSP; P14080; 1YAL.  
 DR MEROPS; C01.UPA; -.  
 DR InterPro; IPR000668; Peptidase\_C1.  
 DR InterPro; IPR000169; SHprot\_acsite.  
 DR Pfam; PF00112; Peptidase\_C1; 1.  
 DR ProDom; PD000158; Peptidase\_C1; 1.  
 DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; PARTIAL.  
 DR PROSITE; PS00640; THIOL\_PROTEASE\_ASN; PARTIAL.  
 DR PROSITE; PS00139; THIOL\_PROTEASE\_CYS; 1.  
 KW Hydrolase; Thiol protease.  
 FT ACT\_SITE 25 25 BY SIMILARITY.  
 FT NON\_TER 43 43  
 SQ SEQUENCE 43 AA; 4685 MW; F4D5945A9386E291 CRC64;

Query Match 9.3%; Score 4; DB 1; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 WRKK 43  
 ||||  
 Db 7 WRKK 10

#### RESULT 6

##### PSAJ\_OENHO

ID PSAJ\_OENHO STANDARD; PRT; 43 AA.  
 AC Q9MTK3;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Photosystem I reaction center subunit IX (PSI-J).  
 GN PSAJ.  
 OS Oenothera hookeri (Hooker's evening primrose).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Myrtales; Onagraceae; Oenothera.  
 OX NCBI\_TaxID=85636;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Johansen;  
 RX MEDLINE=20309318; PubMed=10852478;  
 RA Hupfer H., Swiatek M., Hornung S., Herrmann R.G., Maier R.M.,  
 RA Chiu W.-L., Sears B.;  
 RT "Complete nucleotide sequence of the Oenothera elata plastid  
 RT chromosome, representing plastome I of the five distinguishable  
 RT Eucenothera plastomes.";  
 RL Mol. Gen. Genet. 263:581-585(2000).  
 CC -!- FUNCTION: May help in the organization of the psaE and psaF  
 CC subunits (By similarity).  
 CC -!- SIMILARITY: Belongs to the psaJ family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AJ271079; CAB67177.1; -.  
DR HAMAP; MF\_00522; -; 1.  
DR InterPro; IPR002615; PSI\_PsaJ.  
DR Pfam; PF01701; PSI\_PsaJ; 1.  
DR ProDom; PD004198; PSI\_PsaJ; 1.  
KW Chloroplast; Photosystem I; Photosynthesis; Transmembrane.  
FT TRANSMEM 7 27 POTENTIAL.  
SQ SEQUENCE 43 AA; 4850 MW; EDFC68E6158BA482 CRC64;

Query Match 9.3%; Score 4; DB 1; Length 43;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 LSVA 33  
|||  
Db 8 LSVA 11

#### RESULT 7

##### PSAJ\_ARATH

ID PSAJ\_ARATH STANDARD; PRT; 44 AA.  
AC P56769;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Photosystem I reaction center subunit IX (PSI-J).  
GN PSAJ OR ATCG00630.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=20039611; PubMed=10574454;  
RA Sato S., Nakamura Y., Kaneko T., Asamizu E., Tabata S.;  
RT "Complete structure of the chloroplast genome of Arabidopsis  
RT thaliana.";  
RL DNA Res. 6:283-290(1999).  
CC !- FUNCTION: May help in the organization of the psaE and psaF  
CC subunits (By similarity).  
CC !- SIMILARITY: Belongs to the psaJ family.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP000423; BAA84405.1; -.
DR HAMAP; MF_00522; -; 1.
DR InterPro; IPR002615; PSI_PsaJ.
DR Pfam; PF01701; PSI_PsaJ; 1.
DR ProDom; PD004198; PSI_PsaJ; 1.
KW Chloroplast; Photosystem I; Photosynthesis; Transmembrane.
FT TRANSMEM 7 27 POTENTIAL.
SQ SEQUENCE 44 AA; 5009 MW; B31FFC68FAE6CEBA CRC64;

```

```

Query Match          9.3%; Score 4; DB 1; Length 44;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      30 LSVA 33
        ||||
Db      8 LSVA 11

```

# RESULT 8

## PSAJ\_LOTJA

```

ID PSAJ_LOTJA STANDARD; PRT; 44 AA.
AC Q9BBR3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Photosystem I reaction center subunit IX (PSI-J).
GN PSAJ.
OS Lotus japonicus.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
OX NCBI_TaxID=34305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Accession MG-20;
RX MEDLINE=21082929; PubMed=11214967;
RA Kato T., Kaneko T., Sato S., Nakamura Y., Tabata S.;
RT "Complete structure of the chloroplast genome of a legume, Lotus
RT japonicus.";
RL DNA Res. 7:323-330(2000).
CC -!- FUNCTION: May help in the organization of the psaE and psaF
CC subunits (By similarity).
CC -!- SIMILARITY: Belongs to the psaJ family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP002983; BAB33216.1; -.
DR HAMAP; MF_00522; -; 1.

```

DR InterPro; IPR002615; PSI\_PsaJ.  
 DR Pfam; PF01701; PSI\_PsaJ; 1.  
 DR ProDom; PD004198; PSI\_PsaJ; 1.  
 KW Chloroplast; Photosystem I; Photosynthesis; Transmembrane.  
 FT TRANSMEM 7 27 POTENTIAL.  
 SQ SEQUENCE 44 AA; 5005 MW; 4FAD9A68FAE6C4F5 CRC64;

Query Match 9.3%; Score 4; DB 1; Length 44;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 LSVA 33  
 ||||  
 Db 8 LSVA 11

# RESULT 9

## PSAJ\_ORYSA

ID PSAJ\_ORYSA STANDARD; PRT; 44 AA.  
 AC P12192;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Photosystem I reaction center subunit IX (PSI-J).  
 GN PSAJ.  
 OS Oryza sativa (Rice).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RX MEDLINE=89364698; PubMed=2770692;  
 RA Hiratsuka J., Shimada H., Whittier R., Ishibashi T., Sakamoto M.,  
 RA Mori M., Kondo C., Honji Y., Sun C.-R., Meng B.-Y., Li Y.-Q.,  
 RA Kanno A., Nishizawa Y., Hirai A., Shinozaki K., Sugiura M.;  
 RT "The complete sequence of the rice (Oryza sativa) chloroplast genome:  
 RT intermolecular recombination between distinct tRNA genes accounts for  
 RT a major plastid DNA inversion during the evolution of the cereals."  
 RL Mol. Gen. Genet. 217:185-194(1989).  
 CC -!- FUNCTION: May help in the organization of the psaE and psaF  
 CC subunits (By similarity).  
 CC -!- SIMILARITY: Belongs to the psaJ family.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X15901; CAA33968.1; -.  
 DR PIR; JQ0246; A1RZJ.  
 DR Gramene; P12192; -.

DR HAMAP; MF\_00522; -; 1.  
 DR InterPro; IPR002615; PSI\_PsaJ.  
 DR Pfam; PF01701; PSI\_PsaJ; 1.  
 DR ProDom; PD004198; PSI\_PsaJ; 1.  
 KW Chloroplast; Photosystem I; Photosynthesis; Transmembrane.  
 FT TRANSMEM 7 27 POTENTIAL.  
 SQ SEQUENCE 44 AA; 5050 MW; 4DD9AD29FAF90B85 CRC64;

Query Match 9.3%; Score 4; DB 1; Length 44;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 LSVA 33  
 ||||  
 Db 8 LSVA 11

# RESULT 10

## PSAJ\_SPIOL

ID PSAJ\_SPIOL STANDARD; PRT; 44 AA.  
 AC P17230; Q9M3K9;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Photosystem I reaction center subunit IX (PSI-J).  
 GN PSAJ.  
 OS Spinacia oleracea (Spinach).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
 OX NCBI\_TaxID=3562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Geant d'hiver, and cv. Monatol;  
 RX MEDLINE=21187424; PubMed=11292076;  
 RA Schmitz-Linneweber C., Maier R.M., Alcaraz J.-P., Cottet A.,  
 RA Herrmann R.G., Mache R.;  
 RT "The plastid chromosome of spinach (Spinacia oleracea): complete  
 RT nucleotide sequence and gene organization."  
 RL Plant Mol. Biol. 45:307-315(2001).  
 RN [2]  
 RP SEQUENCE OF 1-16.  
 RX MEDLINE=90242987; PubMed=2185953;  
 RA Ikeuchi M., Hirano A., Hiyama T., Inoue Y.;  
 RT "Polypeptide composition of higher plant photosystem I complex.  
 RT Identification of psaI, psaJ and psaK gene products."  
 RL FEBS Lett. 263:274-278(1990).  
 CC -!- FUNCTION: May help in the organization of the psaE and psaF  
 CC subunits.  
 CC -!- SIMILARITY: Belongs to the psaJ family.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial



CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ400848; CAB88748.1; -.  
 DR HAMAP; MF\_00522; -; 1.  
 DR InterPro; IPR002615; PSI\_PsaJ.  
 DR Pfam; PF01701; PSI\_PsaJ; 1.  
 DR ProDom; PD004198; PSI\_PsaJ; 1.  
 KW Chloroplast; Photosystem I; Photosynthesis; Transmembrane.  
 FT TRANSMEM 7 27 POTENTIAL.  
 SQ SEQUENCE 44 AA; 5077 MW; BB375EEAF244CEBA CRC64;

Query Match 9.3%; Score 4; DB 1; Length 44;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 LSVA 33  
 ||||  
 Db 8 LSVA 11

# RESULT 11

## PSAJ\_TOBAC

ID PSAJ\_TOBAC STANDARD; PRT; 44 AA.  
 AC P12193;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Photosystem I reaction center subunit IX (PSI-J).  
 GN PSAJ.  
 OS Nicotiana tabacum (Common tobacco).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Bright Yellow 4;  
 RA Shinozaki K., Ohme M., Tanaka M., Wakasugi T., Hayashida N.,  
 RA Matsubayashi T., Zaita N., Chunwongse J., Obokata J.,  
 RA Yamaguchi-Shinozaki K., Ohto C., Torazawa K., Meng B.-Y., Sugita M.,  
 RA Deno H., Kamogashira T., Yamada K., Kusuda J., Takaiwa F., Kato A.,  
 RA Tohdoh N., Shimada H., Sugiura M.;  
 RT "The complete nucleotide sequence of the tobacco chloroplast genome:  
 RT its gene organization and expression.";  
 RL EMBO J. 5:2043-2049(1986).  
 CC -!- FUNCTION: May help in the organization of the psaE and psaF  
 CC subunits (By similarity).  
 CC -!- SIMILARITY: Belongs to the psaJ family.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; Z00044; CAA77421.1; -.

DR HAMAP; MF\_00522; -; 1.

DR InterPro; IPR002615; PSI\_PsaJ.

DR Pfam; PF01701; PSI\_PsaJ; 1.

DR ProDom; PD004198; PSI\_PsaJ; 1.

KW Chloroplast; Photosystem I; Photosynthesis; Transmembrane.

FT TRANSMEM 7 27 POTENTIAL.

SQ SEQUENCE 44 AA; 5027 MW; B22DFC68F244CEA4 CRC64;

Query Match 9.3%; Score 4; DB 1; Length 44;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 LSVA 33  
 ||||  
 Db 8 LSVA 11

# RESULT 12

## BD13\_BOVIN

ID BD13\_BOVIN STANDARD; PRT; 42 AA.

AC P46171;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Beta-defensin 13 (BNDB-13) (BNBD-13).

GN DEFB13.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;

OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE.

RC TISSUE=Neutrophils;

RX MEDLINE=93203264; PubMed=8454635;

RA Selsted M.E., Tang Y.Q., Morris W.L., McGuire P.A., Novotny M.J.,

RA Smith W., Henschen A.H., Cullor J.S.;

RT "Purification, primary structures, and antibacterial activities of

RT beta-defensins, a new family of antimicrobial peptides from bovine

RT neutrophils.";

RL J. Biol. Chem. 268:6641-6648(1993).

CC -!- FUNCTION: HAS BACTERICIDAL ACTIVITY. ACTIVE AGAINST E.COLI ML35

CC AND S.AUREUS 502A.

CC -!- TISSUE SPECIFICITY: NEUTROPHILIC GRANULES.

CC -!- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.

DR PIR; D47753; D47753.

DR HSSP; P46170; 1BNB.

DR InterPro; IPR001855; Defensin\_beta.

DR InterPro; IPR006080; Defensin\_mammal.

DR Pfam; PF00711; Defensin\_beta; 1.

DR SMART; SM00048; DEFSN; 1.

KW Antibiotic.

FT DISULFID 9 38 BY SIMILARITY.

FT DISULFID 16 31 BY SIMILARITY.

FT DISULFID 21 39 BY SIMILARITY.  
SQ SEQUENCE 42 AA; 4450 MW; BEA5C5806A04910B CRC64;

Query Match 7.0%; Score 3; DB 1; Length 42;  
Best Local Similarity 100.0%; Pred. No. 1.3e+04;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PMR 3  
|||  
Db 24 PMR 26

#### RESULT 13

##### DLP1\_ORNAN

ID DLP1\_ORNAN STANDARD; PRT; 42 AA.  
AC P82172;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Defensin-like peptide 1 (DLP-1).  
OS Ornithorhynchus anatinus (Duckbill platypus).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.  
OX NCBI\_TaxID=9258;  
RN [1]  
RP SEQUENCE, SYNTHESIS, AND STRUCTURE BY NMR.  
RC TISSUE=Venom, and Venom gland;  
RX MEDLINE=99348045; PubMed=10417345;  
RA Torres A.M., Wang X., Fletcher J.I., Alewood D., Alewood P.F.,  
RA Smith R., Simpson R.J., Nicholson G.M., Sutherland S.K.,  
RA Gallagher C.H., King G.F., Kuchel P.W.;  
RT "Solution structure of a defensin-like peptide from platypus venom."  
RL Biochem. J. 341:785-794(1999).  
CC -!- FUNCTION: Does not show antimicrobial, myotoxic, hemolytic and  
CC cell-promoting activities.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Produced by the venomous spur located on  
CC each male hind leg.  
CC -!- DATABASE: NAME=Protein Spotlight;  
CC NOTE=Issue 29 of December 2002;  
CC WWW="http://www.expasy.org/spotlight/articles/sptlt029.html".  
DR PDB; 1B8W; 15-SEP-99.  
KW Toxin; 3D-structure.  
FT DISULFID 9 39  
FT DISULFID 16 32  
FT DISULFID 24 40  
FT HELIX 10 12  
FT TURN 13 13  
FT STRAND 15 18  
FT TURN 19 20  
FT STRAND 26 26  
FT TURN 28 29  
FT STRAND 37 40  
SQ SEQUENCE 42 AA; 4959 MW; 7ABFF002B8FF1A44 CRC64;

Query Match 7.0%; Score 3; DB 1; Length 42;  
Best Local Similarity 100.0%; Pred. No. 1.3e+04;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 GQK 19  
|||  
Db 36 GQK 38

RESULT 14

GIP\_BOVIN

ID GIP\_BOVIN STANDARD; PRT; 42 AA.  
AC P09680;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE Gastric inhibitory polypeptide (GIP) (Glucose-dependent insulinotropic polypeptide).  
GN GIP.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=85076655; PubMed=6391923;  
RA Carlquist M., Maletti M., Joernvall H., Mutt V.;  
RT "A novel form of gastric inhibitory polypeptide (GIP) isolated from  
RT bovine intestine using a radioreceptor assay. Fragmentation with  
RT staphylococcal protease results in GIP1-3 and GIP4-42, fragmentation  
RT with enterokinase in GIP1-16 and GIP17-42.";  
RL Eur. J. Biochem. 145:573-577(1984).  
CC -!- FUNCTION: POTENT STIMULATOR OF INSULIN SECRETION AND RELATIVELY  
CC POOR INHIBITOR OF GASTRIC ACID SECRETION.  
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.  
DR PIR; S07231; GIBO.  
DR HSSP; P01274; 1GCN.  
DR InterPro; IPR000532; Glucagon.  
DR Pfam; PF00123; hormone2; 1.  
DR SMART; SM00070; GLUCA; 1.  
DR PROSITE; PS00260; GLUCAGON; 1.  
KW Glucagon family; Hormone.  
SQ SEQUENCE 42 AA; 4961 MW; 7DAE3E5C09390F9F CRC64;

Query Match 7.0%; Score 3; DB 1; Length 42;  
Best Local Similarity 100.0%; Pred. No. 1.3e+04;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 AMD 14  
|||  
Db 13 AMD 15

RESULT 15

GIP\_PIG

ID GIP\_PIG STANDARD; PRT; 42 AA.  
AC P01281;  
DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Gastric inhibitory polypeptide (GIP) (Glucose-dependent insulinotropic  
 DE polypeptide).  
 GN GIP.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=81189070; PubMed=7227513;  
 RA Joernvall H., Carlquist M., Kwauk S., Otte S.C., McIntosh C.H.S.,  
 RA Brown J.C., Mutt V.;  
 RT "Amino acid sequence and heterogeneity of gastric inhibitory  
 RT polypeptide (GIP).";  
 RL FEBS Lett. 123:205-210(1981).  
 RN [2]  
 RP SEQUENCE OF 7-42.  
 RC TISSUE=Intestine;  
 RX MEDLINE=93387315; PubMed=8375398;  
 RA Agerberth B., Boman A., Andersson M., Joernvall H., Mutt V.,  
 RA Boman H.G.;  
 RT "Isolation of three antibacterial peptides from pig intestine:  
 RT gastric inhibitory polypeptide (7-42), diazepam-binding inhibitor  
 RT (32-86) and a novel factor, peptide 3910.";  
 RL Eur. J. Biochem. 216:623-629(1993).  
 CC -!- FUNCTION: POTENT STIMULATOR OF INSULIN SECRETION AND RELATIVELY  
 CC POOR INHIBITOR OF GASTRIC ACID SECRETION.  
 CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.  
 DR PIR; A01546; GIPG.  
 DR HSSP; P01274; 1GCN.  
 DR InterPro; IPR000532; Glucagon.  
 DR Pfam; PF00123; hormone2; 1.  
 DR SMART; SM00070; GLUCA; 1.  
 DR PROSITE; PS00260; GLUCAGON; 1.  
 KW Glucagon family; Hormone.  
 FT VARIANT 41 42 MISSING (IN A 2ND COMPONENT).  
 SQ SEQUENCE 42 AA; 4976 MW; 7DAE3E5C093A6F9F CRC64;

Query Match 7.0%; Score 3; DB 1; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+04;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 AMD 14  
 |||  
 Db 13 AMD 15

Search completed: January 7, 2004, 10:26:12  
 Job time : 12 secs

OM protein - protein search, using sw model

Run on: January 7, 2004, 10:27:24 ; Search time 40 Seconds  
(without alignments)  
333.326 Million cell updates/sec

Title: US-09-936-697-6  
Perfect score: 84  
Sequence: 1 QGRSGCSSQSPMRSEIN.....SPTASSQSSATNMAIHRSP 84

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 7743

Minimum DB seq length: 83

Maximum DB seq length: 85

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	% Query		Match Length	DB	ID	Description
	Score	Match				
1	84	100.0	84	21	AAB18942	Peptide derived fr
2	25	29.8	84	21	AAB18938	Peptide derived fr
3	6	7.1	84	22	AAM06662	Human foetal prote
4	5	6.0	83	21	AAG57760	Arabidopsis thalia
5	5	6.0	83	21	AAG58244	Arabidopsis thalia
6	5	6.0	83	22	ABG49367	Human liver peptid
7	5	6.0	83	22	AAU54418	Propionibacterium
8	5	6.0	83	22	ABG11058	Novel human diagno
9	5	6.0	83	22	ABB29372	Peptide #2023 enco
10	5	6.0	83	22	ABB34546	Peptide #2052 enco
11	5	6.0	83	22	ABB19955	Protein #1954 enco
12	5	6.0	83	22	AAM55332	Human brain expres
13	5	6.0	83	22	AAM67729	Human bone marrow
14	5	6.0	83	22	AAM90873	Human immune/haema
15	5	6.0	83	22	AAM15539	Peptide #1973 enco
16	5	6.0	83	22	AAM28029	Peptide #2066 enco
17	5	6.0	83	22	AAM03288	Peptide #1970 enco
18	5	6.0	83	23	ABG37289	Human peptide enco
19	5	6.0	83	23	ABP32408	Human ORF1381 prot
20	5	6.0	83	23	ABB10071	Human cyclin G29.1
21	5	6.0	84	18	AAW28284	Staphylococcus aur
22	5	6.0	84	21	AAG36077	Zea mays protein f
23	5	6.0	84	21	AAG44625	Zea mays protein f
24	5	6.0	84	21	AAG56839	Arabidopsis thalia
25	5	6.0	84	21	AAV94955	Human secreted pro
26	5	6.0	84	22	AAU41307	Propionibacterium
27	5	6.0	84	22	AAU63652	Propionibacterium
28	5	6.0	84	22	AAU64432	Propionibacterium
29	5	6.0	84	22	AAU67297	Propionibacterium
30	5	6.0	84	22	ABG03101	Novel human diagno
31	5	6.0	84	22	AAO13392	Human polypeptide
32	5	6.0	84	22	AAM40586	Human polypeptide
33	5	6.0	84	23	ABP11233	Human ORFX protein
34	5	6.0	85	21	AAG12372	Zea mays protein f
35	5	6.0	85	21	AAG26736	Zea mays protein f
36	5	6.0	85	21	AAG32932	Arabidopsis thalia
37	5	6.0	85	22	AAU53333	Propionibacterium
38	5	6.0	85	22	ABB16594	Human nervous syst
39	5	6.0	85	22	AAM84433	Human immune/haema
40	5	6.0	85	22	AAU17638	Novel signal trans
41	5	6.0	85	22	AAU03385	Human partial fsh2
42	5	6.0	85	23	ABB76265	Escherichia coli t
43	5	6.0	85	23	ABP40669	Staphylococcus epi
44	4	4.8	83	11	AAR05708	Pseudomonas aerugi
45	4	4.8	83	11	AAR07036	Portion of second

# ALIGNMENTS

AAB18942

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PT

XX

XX

XX

Figure 1

Best Local Similarity 100.0%; Pred. No. 3.3e-82;

Db

1 QGRSGCSSQSI SPMRSI SENSLVAMDFSGOKSRVI ENPTEALSVAVEEGLAWRKKGCLRL 60



Qy 61 GTHGSPTASSQSSATNMAIHRSQP 84  
 ||||||||||||||||||  
 Db 61 GTHGSPTASSQSSATNMAIHRSQP 84

RESULT 2

AAB18938

ID AAB18938 standard; peptide; 84 AA.

XX

AC AAB18938;

XX

DT 08-FEB-2001 (first entry)

XX

DE Peptide derived from the PIR or PIR-SH2 domain of Grb7 protein.

XX

KW Phosphorylated insulin receptor interacting region; Grb7 family protein;  
 KW insulin receptor; tyrosine kinase; insulin; insulin-associated disease;  
 KW diabetes; obesity; polycystic ovarian syndrome; syndrome X.

XX

OS Rattus sp.

XX

PN WO200055634-A1.

XX

PD 21-SEP-2000.

XX

PF 14-MAR-2000; 2000WO-FR00613.

XX

PR 15-MAR-1999; 99FR-0003159.

XX

PA (CNRS ) CNRS CENT NAT RECH SCI.

XX

PI Burnol A, Perdereau D, Kasus-Jacobi A, Bereziat V, Girard J;

XX

DR WPI; 2000-587566/55.

XX

PT Fragments of Grb family proteins to identify compounds are useful in  
 PT treating insulin-associated diseases, particularly diabetes and obesity

PT -

XX

PS Claim 2; Page 23-24; 46pp; French.

XX

CC B18937-64 represent the PIR (phosphorylated insulin receptor interacting  
 CC region) or PIR-SH2 (Src homology 2) domains of a Grb7 family protein.  
 CC PIR is the actual binding region but its effect is about 10 times  
 CC greater in presence of SH2 (which by itself is inactive). Agents that  
 CC affect binding between the peptides and the insulin receptor can  
 CC stimulate or inhibit tyrosine kinase activity of the receptor. The  
 CC peptides are used for screening molecules for ability to treat diseases  
 CC in which insulin is implicated. The peptides are used to identify agents  
 CC that are potentially useful for treating insulin-associated diseases,  
 CC particularly diabetes and obesity but also polycystic ovarian syndrome  
 CC and syndrome X.

XX

SQ Sequence 84 AA;

Query Match 29.8%; Score 25; DB 21; Length 84;

Best Local Similarity 100.0%; Pred. No. 6.6e-19;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 NPTEALSVAVEEGLAWRKKGCLRLG 61  
|||||||  
Db 37 NPTEALSVAVEEGLAWRKKGCLRLG 61

RESULT 3

AAM06662

ID AAM06662 standard; Protein; 84 AA.

XX

AC AAM06662;

XX

DT 05-OCT-2001 (first entry)

XX

DE Human foetal protein, SEQ ID NO: 393.

XX

KW Human; foetal protein; cytostatic; immunosuppressive; immunostimulant;  
KW nootropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;  
KW gene therapy; antisense therapy; cancer; immune disorder;  
KW growth disorder; osteoporosis; thrombolytic disorder;  
KW nervous system disorder; inflammation.

XX

OS Homo sapiens.

XX

PN WO200155339-A2.

XX

PD 02-AUG-2001.

XX

PF 25-JAN-2001; 2001WO-US02723.

XX

PR 25-JAN-2000; 2000US-0491404.

PR 15-SEP-2000; 2000US-0663870.

PR 06-NOV-2000; 2000US-0707351.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Yeung G, Ford JE, Boyle BJ, Arterburn MC, Drmanac RA, Tang YT;  
PI Liu C, Asundi V, Zhou P, Werhman T;

XX

DR WPI; 2001-465571/50.

DR N-PSDB; AAH94337.

XX

PT Novel fetal proteins useful for the treatment and diagnosis of diseases  
PT associated with dysfunction of the protein e.g. cancers, immune  
PT disorders, growth disorders, thrombolytic disorders, nervous system  
PT disorders and inflammation -

XX

PS Claim 10; Page 321; 715pp; English.

XX

CC The invention relates to novel foetal polypeptides encoded by  
CC polynucleotides comprising one of 477 sequences fully defined in the  
CC specification. The foetal polynucleotides and polypeptides are  
CC useful in the treatment and diagnosis of diseases such as cancers,  
CC immune disorders, growth disorders (e.g. osteoporosis), thrombolytic  
CC disorders, nervous system disorders and inflammation. The present

CC sequence is a polypeptide encoded by a cDNA assembled using  
CC an expressed sequence tag (EST) found to be expressed in human  
CC foetal tissue cDNA libraries.  
XX  
SQ Sequence 84 AA;

Query Match 7.1%; Score 6; DB 22; Length 84;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 THGSPT 67  
|||  
Db 72 THGSPT 77

RESULT 4

AAG57760

ID AAG57760 standard; Protein; 83 AA.

XX

AC AAG57760;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 74472.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR	06-MAY-1999;	99US-0132487.
PR	07-MAY-1999;	99US-0132863.
PR	11-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.

PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.

PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 6.0%; Score 5; DB 21; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 VAVEE 48  
 ||||  
 Db 68 VAVEE 72

RESULT 5

AAG58244

ID AAG58244 standard; Protein; 83 AA.

XX

AC AAG58244;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 75163.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.



PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.

PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 6.0%; Score 5; DB 21; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 SSQSS 73  
 |||||  
 Db 70 SSQSS 74

# RESULT 6

ABG49367

ID ABG49367 standard; Peptide; 83 AA.

XX

AC ABG49367;

XX

DT 25-FEB-2003 (first entry)

XX

DE Human liver peptide, SEQ ID No 28015.

XX

KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
 KW hypercholesterolaemia; coronary heart disease.

XX

OS Homo sapiens.

XX

PN WO200157273-A2.

XX

PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00664.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488898/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analysing gene expression in human adult liver -  
 XX  
 PS Claim 27; SEQ ID No 28015; 658pp; English.  
 XX  
 CC The invention relates to a single exon nucleic acid probe (SENP) (I) for  
 CC measuring human gene expression in a sample derived from human adult  
 CC liver, comprising one of 13109 defined nucleotide sequences given in the  
 CC specification (or complements/ fragments). The probe hybridises at high  
 CC stringency to a nucleic acid molecule expressed in the human adult  
 CC liver. (I) may be used for predicting, measuring and displaying gene  
 CC expression in samples derived from human adult liver. The genes  
 CC identified may be involved in genetic liver diseases such as cirrhosis,  
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which  
 CC is associated with coronary heart disease. ABG47348-ABG59930 represent  
 CC human liver single exon encoded peptides of the invention.  
 CC Note: The sequence information for this patent does not appear in the  
 CC printed specification but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 83 AA;  
  
 Query Match 6.0%; Score 5; DB 22; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 70 SQSSA 74  
 |||||  
 Db 72 SQSSA 76

RESULT 7

AAU54418

ID AAU54418 standard; Protein; 83 AA.

XX

AC AAU54418;

XX

DT 27-FEB-2002 (first entry)

XX

DE Propionibacterium acnes immunogenic protein #15314.  
 XX  
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.  
 XX  
 OS Propionibacterium acnes.  
 XX  
 PN WO200181581-A2.  
 XX  
 PD 01-NOV-2001.  
 XX  
 PF 20-APR-2001; 2001WO-US12865.  
 XX  
 PR 21-APR-2000; 2000US-199047P.  
 PR 02-JUN-2000; 2000US-208841P.  
 PR 07-JUL-2000; 2000US-216747P.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
 XX  
 DR WPI; 2001-616774/71.  
 DR N-PSDB; AAS59564.  
 XX  
 PT Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris -  
 XX  
 PS Example 1; SEQ ID No 15613; 1069pp; English.  
 XX  
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA).  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO.  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 83 AA;

Query Match

6.0%; Score 5; DB 22; Length 83;

Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 66 PTASS 70  
|||||  
Db 53 PTASS 57

RESULT 8

ABG11058

ID ABG11058 standard; Protein; 83 AA.

XX

AC ABG11058;

XX

DT 18-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #11049.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US08631.

XX

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI; 2001-639362/73.

DR N-PSDB; AAS75245.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -

XX

PS Claim 20; SEQ ID No 41417; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 83 AA;

Query Match 6.0%; Score 5; DB 22; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 73 SATNM 77

||||

Db 44 SATNM 48

#### RESULT 9

ABB29372

ID ABB29372 standard; Peptide; 83 AA.

XX

AC ABB29372;

XX

DT 01-FEB-2002 (first entry)

XX

DE Peptide #2023 encoded by breast cell single exon nucleic acid probe.

XX

KW Human; microarray; single exon probe; gene expression; breast;  
KW disease; cancer.

XX

OS Homo sapiens.

XX

PN WO200157271-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US00662.

XX

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI; 2001-496933/54.

XX

PT New spatially-addressable set of single exon nucleic acid probes,  
PT useful for measuring gene expression in sample derived from human  
PT breast, comprises number of single exon nucleic acid probes -  
XX  
PS Claim 27; SEQ ID NO 12340; 327pp + sequence listing; English.  
XX  
CC The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human breast and BT 474 cells. The method involves contacting  
CC the probes with a collection of detectably labelled nucleic acids  
CC derived from mRNA of human breast, and then measuring the label  
CC bound to each probe of the microarray. The probes are useful for  
CC verifying the expression of regions of genomic DNA predicted to  
CC encode proteins. They are useful for gene discovery, and for  
CC determining predisposition and/or prognosing breast disease. Gene  
CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a peptide encoded by a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 83 AA;

Query Match 6.0%; Score 5; DB 22; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 SQSSA 74  
|||||  
Db 72 SQSSA 76

# RESULT 10

ABB34546

ID ABB34546 standard; Peptide; 83 AA.

XX

AC ABB34546;

XX

DT 04-FEB-2002 (first entry)

XX

DE Peptide #2052 encoded by human foetal liver single exon probe.

XX

KW Human; foetal liver; gene expression; single exon nucleic acid probe.

XX

OS Homo sapiens.

XX

PN WO200157277-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US00669.

XX

PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI; 2001-483447/52.

XX

PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human fetal liver -

XX

PS Claim 27; SEQ ID NO 27181; 639pp + sequence listing; English.

XX

CC The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 83 AA;

Query Match 6.0%; Score 5; DB 22; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 SQSSA 74

|||||

Db 72 SQSSA 76

RESULT 11

ABB19955

ID ABB19955 standard; Protein; 83 AA.

XX

AC ABB19955;

XX

DT 23-JAN-2002 (first entry)

XX

DE Protein #1954 encoded by probe for measuring heart cell gene expression.

XX

KW Human; gene expression; heart; microarray; vascular system;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease.

XX

OS Homo sapiens.

XX

PN WO200157274-A2.



XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00666.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488899/53.  
 XX  
 PT Single exon nucleic acid probes for analyzing gene expression in human  
 PT hearts -  
 XX  
 PS Claim 15; SEQ ID No 21725; 530pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart (see  
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
 CC probe. The probes may be used for predicting, measuring and displaying  
 CC gene expression in samples derived from the human heart via microarrays.  
 CC By measuring gene expression, the probes are useful for predicting,  
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
 CC human heart and vascular system e.g. cardiovascular disease,  
 CC hypertension, cardiac arrhythmias and congenital heart disease.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 83 AA;

Query Match 6.0%; Score 5; DB 22; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 SQSSA 74  
 |||||  
 Db 72 SQSSA 76

# RESULT 12

AAM55332

ID AAM55332 standard; Protein; 83 AA.

XX

AC AAM55332;

XX

DT 05-NOV-2001 (first entry)

XX

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 27437.

XX  
 KW Human; brain expressed exon; gene expression analysis; probe;  
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
 KW epilepsy; cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157275-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00667.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-483446/52.  
 XX  
 PT Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains -  
 XX  
 PS Example 4; SEQ ID NO: 27437; 650pp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is a protein encoded by one of  
 CC the probes of the invention.  
 XX  
 SQ Sequence 83 AA;

Query Match 6.0%; Score 5; DB 22; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 SQSSA 74  
 |||||  
 Db 72 SQSSA 76

RESULT 13  
 AAM67729  
 ID AAM67729 standard; Protein; 83 AA.  
 XX  
 AC AAM67729;  
 XX

DT 06-NOV-2001 (first entry)  
 XX  
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 28035.  
 XX  
 KW Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157276-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00668.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488900/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human bone marrow -  
 XX  
 PS Example 4; SEQ ID NO: 28035; 658pp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention.  
 XX  
 SQ Sequence 83 AA;

Query Match 6.0%; Score 5; DB 22; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 SQSSA 74  
 |||||  
 Db 72 SQSSA 76

#### RESULT 14

AAM90873

ID AAM90873 standard; Protein; 83 AA.

XX

AC AAM90873;

XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen SEQ ID NO:18466.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis.  
XX  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.

PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Barash SC, Ruben SM;

XX

DR WPI; 2001-483426/52.

DR N-PSDB; AAK63654.

XX

PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -

XX

PS Claim 11; SEQ ID NO 18466; 3071pp + Sequence Listing; English.

XX

CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For

CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patients own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 83 AA;

Query Match 6.0%; Score 5; DB 22; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 SSQSS 73  
|||  
Db 20 SSQSS 24

RESULT 15

AAM15539

ID AAM15539 standard; Protein; 83 AA.

XX

AC AAM15539;

XX

DT 12-OCT-2001 (first entry)

XX

DE Peptide #1973 encoded by probe for measuring cervical gene expression.

XX

KW Probe; human; microarray; gene expression; cervical epithelial cell;  
KW cervical cancer.

XX

OS Homo sapiens.

XX

PN WO200157278-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US00670.

XX

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI; 2001-488901/53.

XX

PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human cervical epithelial cells -

XX

PS Claim 27; SEQ ID No 20365; 487pp; English.

XX

CC The present invention relates to human single exon nucleic acid probes  
CC (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded  
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
CC can be used to produce a single exon microarray, which can be used for  
CC measuring human gene expression in a sample derived from human cervical  
CC epithelial cells. By measuring gene expression, the probes are therefore  
CC useful in grading and/or staging of diseases of the cervix, notably  
CC cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 83 AA;

Query Match 6.0%; Score 5; DB 22; Length 83;

Best Local Similarity 100.0%; Pred. No. 1.9e+03;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 SQSSA 74

|||||

Db 72 SQSSA 76

Search completed: January 7, 2004, 10:33:15

Job time : 41 secs



OM protein - protein search, using sw model

Run on: January 7, 2004, 10:32:25 ; Search time 21 Seconds  
(without alignments)  
169.243 Million cell updates/sec

Title: US-09-936-697-6  
Perfect score: 84  
Sequence: 1 QGRSGCSSQSPMRSEN.....SPTASSQSSATNMAIHRSQP 84

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 915

Minimum DB seq length: 83

Maximum DB seq length: 85

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	6	7.1	83	4	US-09-252-991A-21499	Sequence 21499, A
2	5	6.0	83	4	US-09-328-352-6909	Sequence 6909, Ap
3	5	6.0	85	4	US-09-162-564-7	Sequence 7, Appli
4	5	6.0	85	4	US-09-134-001C-5514	Sequence 5514, Ap
5	4	4.8	83	3	US-08-851-843A-9	Sequence 9, Appli
6	4	4.8	83	3	US-09-382-155-11	Sequence 11, Appl
7	4	4.8	83	3	US-08-974-549A-191	Sequence 191, App
8	4	4.8	83	3	US-09-074-044A-11	Sequence 11, Appl
9	4	4.8	83	3	US-08-854-050-9	Sequence 9, Appli
10	4	4.8	83	3	US-09-357-251-4	Sequence 4, Appli
11	4	4.8	83	4	US-09-430-323-9	Sequence 9, Appli

12	4	4.8	83	4	US-09-222-575-62	Sequence 62, Appl
13	4	4.8	83	4	US-09-389-681-62	Sequence 62, Appl
14	4	4.8	83	4	US-09-620-405B-62	Sequence 62, Appl
15	4	4.8	83	4	US-09-482-273-112	Sequence 112, App
16	4	4.8	83	4	US-09-252-991A-17617	Sequence 17617, A
17	4	4.8	83	4	US-09-252-991A-18845	Sequence 18845, A
18	4	4.8	83	4	US-09-198-452A-969	Sequence 969, App
19	4	4.8	83	4	US-09-328-352-5531	Sequence 5531, Ap
20	4	4.8	83	4	US-09-732-210-1197	Sequence 1197, Ap
21	4	4.8	83	4	US-09-339-338-62	Sequence 62, Appl
22	4	4.8	83	4	US-09-433-826B-62	Sequence 62, Appl
23	4	4.8	83	4	US-09-604-287A-62	Sequence 62, Appl
24	4	4.8	84	2	US-08-744-670-5	Sequence 5, Appli
25	4	4.8	84	2	US-09-149-933-5	Sequence 5, Appli
26	4	4.8	84	3	US-08-648-322-10	Sequence 10, Appl
27	4	4.8	84	3	US-09-404-671-6	Sequence 6, Appli
28	4	4.8	84	3	US-08-905-223-404	Sequence 404, App
29	4	4.8	84	3	US-08-431-517F-9	Sequence 9, Appli
30	4	4.8	84	4	US-09-439-313-571	Sequence 571, App
31	4	4.8	84	4	US-09-227-357-597	Sequence 597, App
32	4	4.8	84	4	US-08-857-389-10	Sequence 10, Appl
33	4	4.8	84	4	US-09-205-258-395	Sequence 395, App
34	4	4.8	84	4	US-09-252-991A-18152	Sequence 18152, A
35	4	4.8	84	4	US-09-252-991A-18299	Sequence 18299, A
36	4	4.8	84	4	US-09-252-991A-18487	Sequence 18487, A
37	4	4.8	84	4	US-09-252-991A-22043	Sequence 22043, A
38	4	4.8	84	4	US-09-252-991A-27762	Sequence 27762, A
39	4	4.8	84	4	US-09-198-452A-452	Sequence 452, App
40	4	4.8	84	4	US-09-328-352-7808	Sequence 7808, Ap
41	4	4.8	84	4	US-09-107-532A-3736	Sequence 3736, Ap
42	4	4.8	84	4	US-09-107-532A-6677	Sequence 6677, Ap
43	4	4.8	84	6	5171684-3	Patent No. 5171684
44	4	4.8	84	6	5180813-4	Patent No. 5180813
45	4	4.8	85	1	US-08-615-170-24	Sequence 24, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-252-991A-21499

; Sequence 21499, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 21499

; LENGTH: 83  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-21499

Query Match 7.1%; Score 6; DB 4; Length 83;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 SPTASS 70  
|||||  
Db 2 SPTASS 7

RESULT 2

US-09-328-352-6909  
; Sequence 6909, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 6909  
; LENGTH: 83  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-6909

Query Match 6.0%; Score 5; DB 4; Length 83;  
Best Local Similarity 100.0%; Pred. No. 5.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 ENSLV 23  
|||||  
Db 49 ENSLV 53

RESULT 3

US-09-162-564-7  
; Sequence 7, Application US/09162564  
; Patent No. 6379664  
; GENERAL INFORMATION:  
; APPLICANT: Lou, Marjorie  
; APPLICANT: Raghavachari, Nalini  
; APPLICANT: Qiao, Fengyu  
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR THE PREVENTION AND TREATMENT  
; TITLE OF INVENTION: OF OXIDATIVE DAMAGE IN OCULAR TISSUES  
; FILE REFERENCE: UNL-98-2-1  
; CURRENT APPLICATION NUMBER: US/09/162,564  
; CURRENT FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 7  
; LENGTH: 85  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-09-162-564-7

Query Match 6.0%; Score 5; DB 4; Length 85;  
Best Local Similarity 100.0%; Pred. No. 5.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRSGC 6  
|||||  
Db 7 GRSGC 11

RESULT 4

US-09-134-001C-5514  
; Sequence 5514, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 5514  
; LENGTH: 85  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-5514

Query Match 6.0%; Score 5; DB 4; Length 85;  
Best Local Similarity 100.0%; Pred. No. 5.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 SSQSI 11  
|||||  
Db 81 SSQSI 85

RESULT 5

US-08-851-843A-9  
; Sequence 9, Application US/08851843A  
; Patent No. 6093809  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.

; APPLICANT: Harley, Calvin  
 ; APPLICANT: Andrews, William H.  
 ; TITLE OF INVENTION: No. 6093809e1 Telomerase  
 ; NUMBER OF SEQUENCES: 225  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Townsend and Townsend and Crew LLP  
 ; STREET: Two Embarcadero Center, 8th Floor  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: United States of America  
 ; ZIP: 94111  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/851,843A  
 ; FILING DATE: 06-MAY-1997  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/846,017  
 ; FILING DATE: 25-APR-1997  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/844,419  
 ; FILING DATE: 18-APR-1997  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/724,643  
 ; FILING DATE: 01-OCT-1996  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Apple, Randolph T.  
 ; REGISTRATION NUMBER: 36,429  
 ; REFERENCE/DOCKET NUMBER: 015389-002930US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 576-0200  
 ; TELEFAX: (415) 576-0300  
 ; INFORMATION FOR SEQ ID NO: 9:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 83 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: not relevant  
 ; TOPOLOGY: not relevant  
 ; MOLECULE TYPE: peptide  
 US-08-851-843A-9

Query Match 4.8%; Score 4; DB 3; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 5.9e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 54 KKGC 57  
 |||  
 Db 25 KKGC 28

RESULT 6

US-09-382-155-11

```
; Sequence 11, Application US/09382155B
; Patent No. 6160095
; GENERAL INFORMATION:
; APPLICANT: CHAUDHARY, PREET M
; APPLICANT: HOOD, LEROY
; TITLE OF INVENTION: PROTEINS CAPABLE OF REGULATING NF-kB, JNK AND APOPTOSIS
; TITLE OF INVENTION: PATHWAYS AND METHODS OF USING THE SAME
; FILE REFERENCE: Chaudhary
; CURRENT APPLICATION NUMBER: US/09/382,155B
; CURRENT FILING DATE: 1999-08-24
; EARLIER APPLICATION NUMBER: 09/074,044
; EARLIER FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 83
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-382-155-11
```

```
Query Match          4.8%; Score 4; DB 3; Length 83;
Best Local Similarity 100.0%; Pred. No. 5.9e+03;
Matches      4; Conservative    0; Mismatches    0; Indels      0; Gaps      0;
```

Qy 25 MDFS 28

||||

Db 1 MDFS 4

RESULT 7

US-08-974-549A-191

```
; Sequence 191, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim.
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 191:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 83 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-549A-191

```

```

Query Match          4.8%; Score 4; DB 3; Length 83;
Best Local Similarity 100.0%; Pred. No. 5.9e+03;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      54 KKGC 57
        ||||
Db      25 KKGC 28

```

RESULT 8

US-09-074-044A-11

; Sequence 11, Application US/09074044A

; Patent No. 6207458

; GENERAL INFORMATION:

; APPLICANT: CHAUDHARY, PREET M

; APPLICANT: HOOD, LEROY

; TITLE OF INVENTION: PROTEINS CAPABLE OF REGULATING NK-kB, JNK AND

; TITLE OF INVENTION: APOPTOSIS PATHWAYS AND METHODS OF USING THE SAME

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS

; STREET: 2405 GRAND BLVD., SUITE 400

; CITY: KANSAS CITY

; STATE: MISSOURI

; COUNTRY: USA

; ZIP: 64108

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/074,044A

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: COLLINS, JOHN M

; REGISTRATION NUMBER: 26,262

; REFERENCE/DOCKET NUMBER: 26588

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 816/474-9050

; TELEFAX: 816/474-9057

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 83 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: not relevant

; MOLECULE TYPE: protein

; FRAGMENT TYPE: N-terminal

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

US-09-074-044A-11

Query Match 4.8%; Score 4; DB 3; Length 83;

Best Local Similarity 100.0%; Pred. No. 5.9e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 MDFS 28

||||

Db 1 MDFS 4

RESULT 9



US-08-854-050-9  
; Sequence 9, Application US/08854050  
; Patent No. 6261836  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: No. 6261836el Telomerase  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/854,050  
; FILING DATE: 09-MAY-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-002930US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 83 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant

; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
US-08-854-050-9

Query Match 4.8%; Score 4; DB 3; Length 83;  
Best Local Similarity 100.0%; Pred. No. 5.9e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 54 KKG C 57  
|||  
Db 25 KKG C 28

RESULT 10

US-09-357-251-4

; Sequence 4, Application US/09357251  
; Patent No. 6271441  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Layo O.  
; APPLICANT: Orozco, Buddy  
; APPLICANT: Schwaber, James S.  
; TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase  
; FILE REFERENCE: BB-1193  
; CURRENT APPLICATION NUMBER: US/09/357,251  
; CURRENT FILING DATE: 1999-07-20  
; EARLIER APPLICATION NUMBER: 60/093,530  
; EARLIER FILING DATE: July 21, 1998  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 4  
; LENGTH: 83  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (76)  
US-09-357-251-4

Query Match 4.8%; Score 4; DB 3; Length 83;  
Best Local Similarity 100.0%; Pred. No. 5.9e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 63 HGSP 66  
|||  
Db 80 HGSP 83

RESULT 11

US-09-430-323-9

; Sequence 9, Application US/09430323  
; Patent No. 6309867  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; Lingner, Joachim  
; Nakamura, Toru  
; Chapman, Karen B.

```

; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 83 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-430-323-9

```

```

Query Match          4.8%; Score 4; DB 4; Length 83;
Best Local Similarity 100.0%; Pred. No. 5.9e+03;
Matches    4; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

```

```

Qy      54 KKGC 57
        ||||
Db      25 KKGC 28

```

RESULT 12

US-09-222-575-62

; Sequence 62, Application US/09222575  
 ; Patent No. 6387697  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yuqiu, Jiang  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Xu, Jiangchun  
 ; TITLE OF INVENTION: Compositions for the Treatment and Diagnosis of Breast Cancer  
 ; TITLE OF INVENTION: and Methods for Their Use  
 ; FILE REFERENCE: 210121.470  
 ; CURRENT APPLICATION NUMBER: US/09/222,575  
 ; CURRENT FILING DATE: 1998-12-28  
 ; NUMBER OF SEQ ID NOS: 174  
 ; SOFTWARE: FastSEQ for Windows Version 3.0  
 ; SEQ ID NO 62  
 ; LENGTH: 83  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 US-09-222-575-62

Query Match 4.8%; Score 4; DB 4; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 5.9e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 RSIS 18  
 ||||  
 Db 28 RSIS 31

RESULT 13

US-09-389-681-62

; Sequence 62, Application US/09389681A  
 ; Patent No. 6518237  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yuqiu, Jiang  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Xu, Jiangchun  
 ; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE  
 ; FILE REFERENCE: 210121.470C3  
 ; CURRENT APPLICATION NUMBER: US/09/389,681A  
 ; CURRENT FILING DATE: 1999-09-02  
 ; NUMBER OF SEQ ID NOS: 463  
 ; SOFTWARE: FastSEQ for Windows Version 3.0  
 ; SEQ ID NO 62  
 ; LENGTH: 83  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-09-389-681-62

Query Match 4.8%; Score 4; DB 4; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 5.9e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 RSIS 18  
|||  
Db 28 RSIS 31

RESULT 14

US-09-620-405B-62

; Sequence 62, Application US/09620405B

; Patent No. 6528054

; GENERAL INFORMATION:

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Xu, Jiangchun

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Hepler, William T.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER

; FILE REFERENCE: 210121.470C8

; CURRENT APPLICATION NUMBER: US/09/620,405B

; CURRENT FILING DATE: 2000-07-20

; NUMBER OF SEQ ID NOS: 495

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 62

; LENGTH: 83

; TYPE: PRT

; ORGANISM: Homo sapien

US-09-620-405B-62

Query Match 4.8%; Score 4; DB 4; Length 83;

Best Local Similarity 100.0%; Pred. No. 5.9e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 RSIS 18  
|||  
Db 28 RSIS 31

RESULT 15

US-09-482-273-112

; Sequence 112, Application US/09482273

; Patent No. 6534631

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 71 Human Secreted Proteins

; FILE REFERENCE: PZ030P1

; CURRENT APPLICATION NUMBER: US/09/482,273

; CURRENT FILING DATE: 2000-01-13

; EARLIER APPLICATION NUMBER: PCT/US99/15849

; EARLIER FILING DATE: 1999-07-14

; EARLIER APPLICATION NUMBER: 60/092,921

; EARLIER FILING DATE: 1998-07-15

; EARLIER APPLICATION NUMBER: 60/092,922

; EARLIER FILING DATE: 1998-07-15

; EARLIER APPLICATION NUMBER: 60/092,956

; EARLIER FILING DATE: 1998-07-15  
; NUMBER OF SEQ ID NOS: 267  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 112  
; LENGTH: 83  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (83)  
; OTHER INFORMATION: Xaa equals stop translation  
US-09-482-273-112

Query Match 4.8%; Score 4; DB 4; Length 83;  
Best Local Similarity 100.0%; Pred. No. 5.9e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 LRLG 61  
|||  
Db 46 LRLG 49

Search completed: January 7, 2004, 10:35:33  
Job time : 22 secs

OM protein - protein search, using sw model

Run on: January 7, 2004, 10:31:49 ; Search time 21 Seconds  
(without alignments)  
384.675 Million cell updates/sec

Title: US-09-936-697-6  
Perfect score: 84  
Sequence: 1 QGRSGCSSQSPMRISSEN.....SPTASSQSSATNM̄IHRSQP 84

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1543

Minimum DB seq length: 83

Maximum DB seq length: 85

Post-processing: Listing first 45 summaries

Database : PIR\_76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	5	6.0	83	2 T30628	hypothetical prote
2	5	6.0	84	2 S72700	probable ketoacyl
3	5	6.0	85	1 GDEC	glutaredoxin 1 - E
4	5	6.0	85	2 C25974	39K rod phycocyani
5	4	4.8	83	2 B84679	hypothetical prote
6	4	4.8	83	2 S07941	proteinase inhibit
7	4	4.8	83	2 S34151	Ig mu chain - axol
8	4	4.8	83	2 S34095	Ig kappa chain V r
9	4	4.8	83	2 S24211	Ig kappa chain V r
10	4	4.8	83	2 S42065	histone H3.1 - Lep
11	4	4.8	83	2 S73226	ribosomal protein
12	4	4.8	83	2 F81267	30S ribosomal prot
13	4	4.8	83	2 S61086	M protein precurs

14	4	4.8	83	2	S41851	coat protein - Phy
15	4	4.8	83	2	A82608	plasmid stabilizat
16	4	4.8	83	2	T17752	hypothetical prote
17	4	4.8	83	2	A64644	hypothetical prote
18	4	4.8	83	2	E83512	hypothetical prote
19	4	4.8	83	2	A33854	outer membrane lip
20	4	4.8	83	2	G81806	hypothetical prote
21	4	4.8	83	2	H81084	hypothetical prote
22	4	4.8	83	2	A82386	hypothetical prote
23	4	4.8	83	2	F75604	hypothetical prote
24	4	4.8	83	2	S72951	hypothetical prote
25	4	4.8	83	2	T35979	hypothetical prote
26	4	4.8	83	2	T36350	probable regulator
27	4	4.8	83	2	S77792	hypothetical prote
28	4	4.8	83	2	F82891	hypothetical prote
29	4	4.8	83	2	E82887	hypothetical prote
30	4	4.8	83	2	H69183	hypothetical prote
31	4	4.8	83	2	T48068	hypothetical prote
32	4	4.8	83	2	T03201	hypothetical prote
33	4	4.8	83	2	S46930	teg292 protein - m
34	4	4.8	83	2	E82744	ferrous iron trans
35	4	4.8	83	2	A97775	hypothetical prote
36	4	4.8	83	2	B95974	probable transcrip
37	4	4.8	83	2	B95352	protein [imported
38	4	4.8	83	2	F89920	conserved hypothet
39	4	4.8	84	1	JU0152	acrosin/trypsin in
40	4	4.8	84	2	T14543	peroxidase (EC 1.1
41	4	4.8	84	2	A53012	serum response fac
42	4	4.8	84	2	JX0332	tenecin 1 precursor
43	4	4.8	84	2	E28840	Ig kappa chain V r
44	4	4.8	84	2	D97284	ribosomal protein
45	4	4.8	84	2	T03787	metallothionein-li

#### ALIGNMENTS

##### RESULT 1

T30628

hypothetical protein 26L - Molluscum contagiosum virus 1

N;Alternate names: MC026L

C;Species: Molluscum contagiosum virus 1

C;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 11-May-2000

C;Accession: T30628

R;Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.  
Science 273, 813-816, 1996

A;Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific  
host response-evasion genes.

A;Reference number: Z20876; MUID:96325459; PMID:8670425

A;Accession: T30628

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-83 <SEN>

A;Cross-references: EMBL:U60315; PIDN:AAC55154.1

C;Genetics:

A;Note: MC026L



Query Match 6.0%; Score 5; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 6.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 55 KGCLR 59  
|||||  
Db 32 KGCLR 36

RESULT 2

S72700

probable ketoacyl synthase kas - Mycobacterium leprae

N;Alternate names: Lepb1170\_C1\_189 protein

C;Species: Mycobacterium leprae

C;Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 22-Oct-1999

C;Accession: S72700

R;Smith, D.R.; Robison, K.

submitted to the EMBL Data Library, November 1993

A;Description: Mycobacterium leprae cosmid B1170.

A;Reference number: S72693

A;Accession: S72700

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-84 <SMI>

A;Cross-references: EMBL:U00010; NID:g466780; PIDN:AAA17064.1; PID:g466788

C;Genetics:

A;Gene: kas

A;Start codon: GTG

Query Match 6.0%; Score 5; DB 2; Length 84;  
Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 SSQSS 73  
|||||  
Db 16 SSQSS 20

RESULT 3

GDEC

glutaredoxin 1 - Escherichia coli (strain K-12)

N;Alternate names: thioltransferase

C;Species: Escherichia coli

C;Date: 19-Feb-1984 #sequence\_revision 19-Feb-1984 #text\_change 01-Mar-2002

C;Accession: A00283; A24397; I59418; A64823; A39568

R;Hoeog, J.O.; Joernvall, H.; Holmgren, A.; Carlquist, M.; Persson, M.

Eur. J. Biochem. 136, 223-232, 1983

A;Title: The primary structure of Escherichia coli glutaredoxin. Distant  
homology with thioredoxins in a superfamily of small proteins with a redox-  
active cystine disulfide/cysteine dithiol.

A;Reference number: A00283; MUID:84004402; PMID:6352262

A;Accession: A00283

A;Molecule type: protein

A;Residues: 1-85 <HO1>

A;Experimental source: K-12, strain C10-17

R;Hoeog, J.O.; von Bahr-Lindstroem, H.; Joernvall, H.; Holmgren, A.

Gene 43, 13-21, 1986

A;Title: Cloning and expression of the glutaredoxin (grx) gene of *Escherichia coli*.  
 A;Reference number: A24397; MUID:87005940; PMID:3530878  
 A;Accession: A24397  
 A;Molecule type: DNA  
 A;Residues: 1-85 <HO2>  
 A;Cross-references: GB:M13449; NID:g146272; PIDN:AAA23936.1; PID:g146273  
 R;Chatterjee, P.K.; Sternberg, N.L.  
 Proc. Natl. Acad. Sci. U.S.A. 92, 8950-8954, 1995  
 A;Title: A general genetic approach in *Escherichia coli* for determining the mechanism(s) of action of tumoricidal agents: application to DMP 840, a tumoricidal agent.  
 A;Reference number: I59418; MUID:96004656; PMID:7568050  
 A;Accession: I59418  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-85 <RES>  
 A;Cross-references: EMBL:U18655; NID:g609323; PIDN:AAC43449.1; PID:g609325  
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.; Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.; Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A;Title: The complete genome sequence of *Escherichia coli* K-12.  
 A;Reference number: A64720; MUID:97426617; PMID:9278503  
 A;Accession: A64823  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-85 <BLAT>  
 A;Cross-references: GB:AE000187; GB:U00096; NID:g1787070; PIDN:AAC73936.1; PID:g1787073; UWGP:b0849  
 A;Experimental source: strain K-12, substrain MG1655  
 R;Sandberg, V.A.; Kren, B.; Fuchs, J.A.; Woodward, C.  
 Biochemistry 30, 5475-5484, 1991  
 A;Title: *Escherichia coli* glutaredoxin: cloning and overexpression, thermodynamic stability of the oxidized and reduced forms, and report of an N-terminal extended species.  
 A;Reference number: A39568; MUID:91242463; PMID:2036416  
 A;Accession: A39568  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 'MRREI',1-15 <SAN>  
 C;Genetics:  
 A;Gene: grxA; grx  
 A;Map position: 19 min  
 C;Function:  
 A;Description: the disulfide bond functions as an electron carrier in the glutathione-dependent synthesis of deoxyribonucleotides from ribonucleotides by the enzyme ribonucleotide reductase; in addition, it is also involved in reducing some disulfides in a coupled system with glutathione reductase  
 A;Pathway: deoxyribonucleotide biosynthesis  
 C;Superfamily: glutaredoxin; glutaredoxin homology  
 C;Keywords: deoxyribonucleotide biosynthesis; electron transfer; monomer; redox-active disulfide  
 F;1-85/Domain: glutaredoxin homology <GLUT>  
 F;11-14/Disulfide bonds: redox-active #status experimental

Query Match

6.0%; Score 5; DB 1; Length 85;

Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRS GC 6  
|||||  
Db 7 GRS GC 11

RESULT 4

C25974

39K rod phycocyanin linker protein - *Calothrix* sp.

C;Species: *Calothrix* sp.

C;Date: 16-Aug-1988 #sequence\_revision 16-Aug-1988 #text\_change 15-Oct-1999

C;Accession: C25974

R;Lomax, T.L.; Conley, P.B.; Schilling, J.; Grossman, A.R.

J. Bacteriol. 169, 2675-2684, 1987

A;Title: Isolation and characterization of light-regulated phycobilisome linker polypeptide genes and their transcription as a polycistronic mRNA.

A;Reference number: A25974; MUID:87222193; PMID:3108238

A;Note: *Fremyella diplosiphon*

A;Accession: C25974

A;Molecule type: DNA

A;Residues: 1-85 <LOM>

A;Cross-references: GB:M16490; NID:g148557; PIDN:AAA24888.1; PID:g148561

A;Experimental source: UTEX 481

Query Match 6.0%; Score 5; DB 2; Length 85;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 ENPTE 40  
|||||  
Db 76 ENPTE 80

RESULT 5

B84679

hypothetical protein At2g27970 [imported] - *Arabidopsis thaliana*

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 17-May-2002

C;Accession: B84679

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.;

Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblyum, T.V.; Buell,

C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.; Cronin,

L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.; Adams,

M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H.M.; Somerville, C.R.; Copenhaver,

G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser,

C.M.; Venter, J.C.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: B84679

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-83 <STO>

A;Cross-references: GB:AE002093; NID:g4510419; PIDN:AAD21505.1; GSPDB:GN00139

C;Genetics:  
A;Gene: At2g27970  
A;Map position: 2  
C;Superfamily: cell division control protein CKS1

Query Match 4.8%; Score 4; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 7.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 AIHR 81  
|||  
Db 56 AIHR 59

#### RESULT 6

S07941

proteinase inhibitor IV - soybean

C;Species: Glycine max (soybean)

C;Date: 02-Dec-1993 #sequence\_revision 01-Dec-1995 #text\_change 18-Aug-2000

C;Accession: S07941

R;Joudrier, P.E.; Foard, D.E.; Floener, L.A.; Larkins, B.A.

Plant Mol. Biol. 10, 35-42, 1987

A;Title: Isolation and sequence of cDNA encoding the soybean protease inhibitors  
PI IV and C-II.

A;Reference number: S07405

A;Accession: S07941

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-83 <JOU>

A;Cross-references: EMBL:M20733; NID:g169946; PIDN:AAA33954.1; PID:g169947

C;Superfamily: Bowman-Birk proteinase inhibitor; Bowman-Birk inhibitor repeat  
homology

F;25-51/Domain: Bowman-Birk inhibitor repeat homology <BB1>

F;52-77/Domain: Bowman-Birk inhibitor repeat homology <BBI2>

Query Match 4.8%; Score 4; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 7.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 RSQP 84  
|||  
Db 59 RSQP 62

#### RESULT 7

S34151

Ig mu chain - axolotl (fragment)

C;Species: Ambystoma mexicanum (axolotl)

C;Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 08-Sep-1997

C;Accession: S34151

R;Fellah, J.S.; Jacques, C.; Charlemagne, J.

submitted to the EMBL Data Library, June 1993

A;Reference number: S34149

A;Accession: S34151

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-83 <FEL>

A;Cross-references: EMBL:X73555; NID:g313180; PID:g313181  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: immunoglobulin

Query Match 4.8%; Score 4; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 7.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 KSRV 34  
|||  
Db 27 KSRV 30

RESULT 8

S34095

Ig kappa chain V region - human

C;Species: Homo sapiens (man)

C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 21-Jan-2000

C;Accession: S34095

R;Wagner, S.D.; Luzzatto, L.

Eur. J. Immunol. 23, 391-397, 1993

A;Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distributed over a large portion of the V-kappa locus and do not show somatic mutation.

A;Reference number: S34076; MUID:93170387; PMID:8436174

A;Accession: S34095

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-83 <WAG>

A;Cross-references: EMBL:X67179

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;1-80/Domain: immunoglobulin homology <IMM>

Query Match 4.8%; Score 4; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 7.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 SSQS 10  
|||  
Db 10 SSQS 13

RESULT 9

S24211

Ig kappa chain V region (V3a) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 19-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 23-Jul-1999

C;Accession: S24211

R;Pargent, W.; Meindl, A.; Thiebe, R.; Mitzel, S.; Zachau, H.G.

Eur. J. Immunol. 21, 1821-1827, 1991

A;Title: The human immunoglobulin kappa locus. Characterization of the duplicated O regions.

A;Reference number: S24205; MUID:91330953; PMID:1907917

A;Accession: S24211

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-83 <PAR>  
A;Cross-references: EMBL:X59317; NID:g33270; PIDN:CAA42004.1; PID:g33271  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin

Query Match 4.8%; Score 4; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 7.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 SSQS 10  
|||  
Db 7 SSQS 10

#### RESULT 10

S42065

histone H3.1 - Leptothorax acervorum (fragment)

C;Species: Leptothorax acervorum

C;Date: 19-Mar-1997 #sequence\_revision 18-Jul-1997 #text\_change 23-Jul-1999

C;Accession: S42065

R;Baur, A.; Stetzer, N.E.; Buschinger, A.; Zimmermann, F.K.

submitted to the EMBL Data Library, February 1994

A;Description: Cloning of two differentially expressed reverse transcription fragments of the histone 3 gene of leptothorax acervorum (Hymenoptera Formicidae).

A;Reference number: S42065

A;Accession: S42065

A;Molecule type: mRNA

A;Residues: 1-83 <BAU>

A;Cross-references: EMBL:X77742; NID:g456194; PIDN:CAA54792.1; PID:g456195

C;Genetics:

A;Gene: H3.1

C;Superfamily: histone H3

C;Keywords: chromosomal protein; DNA binding; nucleosome core; nucleus

Query Match 4.8%; Score 4; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 7.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 71 QSSA 74  
|||  
Db 53 QSSA 56

#### RESULT 11

S73226

ribosomal protein S17, chloroplast - red alga (Porphyra purpurea) chloroplast

C;Species: chloroplast Porphyra purpurea

C;Date: 19-Mar-1997 #sequence\_revision 09-May-1997 #text\_change 13-Aug-1999

C;Accession: S73226

R;Reith, M.; Munholland, J.

Plant Mol. Biol. Rep. 13, 333-335, 1995

A;Title: Complete nucleotide sequence of the Porphyra purpurea chloroplast genome.

A;Reference number: S73108

A;Accession: S73226

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA  
 A;Residues: 1-83 <REI>  
 A;Cross-references: EMBL:U38804; NID:g1276652; PIDN:AAC08191.1; PID:g1276771  
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995  
 C;Genetics:  
 A;Gene: rps17  
 A;Genome: chloroplast  
 C;Superfamily: Escherichia coli ribosomal protein S17  
 C;Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 4.8%; Score 4; DB 2; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 7.6e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 VAVE 47  
 ||||  
 Db 21 VAVE 24

# RESULT 12

F81267  
 30S ribosomal protein S17 Cj1698c [imported] - Campylobacter jejuni (strain NCTC 11168)

C;Species: Campylobacter jejuni  
 C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
 C;Accession: F81267  
 R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Davies, R.M.; Feltwell, T.; Holroyd, S.; Jagels, K.; Karlyshev, A.; Moule, S.; Pallen, M.J.; Penn, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell, B.G.  
 Nature 403, 665-668, 2000

A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences.

A;Reference number: A81250; MUID:20150912; PMID:10688204

A;Accession: F81267

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-83 <PAR>

A;Cross-references: GB:AL139079; GB:AL111168; NID:g6968971; PIDN:CAB73684.1; PID:g6969113; GSPDB:GN00120; CJSP:Cj1698c

A;Experimental source: serotype O2, strain NCTC 11168

C;Genetics:

A;Gene: rpsQ; Cj1698c

C;Superfamily: Escherichia coli ribosomal protein S17

Query Match 4.8%; Score 4; DB 2; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 7.6e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 VAVE 47  
 ||||  
 Db 59 VAVE 62

# RESULT 13

S61086

M protein precursor - Streptococcus pyogenes (serotype M68) (fragment)  
 C;Species: Streptococcus pyogenes  
 A;Variety: serotype M68  
 C;Date: 15-Feb-1996 #sequence\_revision 19-Apr-1996 #text\_change 16-Feb-1997  
 C;Accession: S61086; S60825  
 R;Whatmore, A.; Kapur, V.; Sullivan, D.; Musser, J.; Kehoe, M.  
 submitted to the EMBL Data Library, July 1994  
 A;Description: Noncongruent relationships between variation in emm1 gene  
 sequences and the population genetic structure of group A Streptococci.  
 A;Reference number: S61072  
 A;Accession: S61086  
 A;Molecule type: DNA  
 A;Residues: 1-83 <WHA>  
 A;Cross-references: EMBL:U11997  
 R;Whatmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.  
 Mol. Microbiol. 14, 619-631, 1994  
 A;Title: Non-congruent relationships between variation in emm gene sequences and  
 the population genetic structure of group A streptococci.  
 A;Reference number: S60784; MUID:95198537; PMID:7891551  
 A;Accession: S60825  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: DNA  
 A;Residues: 18-43, 'A', 45-78 <WHW>  
 A;Cross-references: EMBL:U11997  
 C;Genetics:  
 A;Gene: emm68  
 C;Superfamily: M5 protein  
 F;1-30/Domain: signal sequence (fragment) #status predicted <SIG>  
 F;31-83/Product: M protein (fragment) #status predicted <MAT>

Query Match 4.8%; Score 4; DB 2; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 7.6e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 SVAV 46  
 ||||  
 Db 10 SVAV 13

#### RESULT 14

S41851  
 coat protein - Physalis mottle virus (fragment)  
 C;Species: Physalis mottle virus  
 C;Date: 19-Mar-1997 #sequence\_revision 17-Sep-1997 #text\_change 11-May-2000  
 C;Accession: S41851  
 R;Peter, R.; Peter, C.; Dupin, A.; Witz, J.  
 C. R. Acad. Sci. III 309, 599-604, 1989  
 A;Title: A problem of tymovirus taxonomy: comparison of the coat proteins of two  
 Belladonna mottle viruses.  
 A;Reference number: S41850; MUID:90058208; PMID:2510912  
 A;Accession: S41851  
 A;Molecule type: protein  
 A;Residues: 1-83 <PET>  
 C;Superfamily: turnip yellow mosaic virus coat protein  
 C;Keywords: coat protein

Query Match 4.8%; Score 4; DB 2; Length 83;



Best Local Similarity 100.0%; Pred. No. 7.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 66 PTAS 69  
||||  
Db 63 PTAS 66

RESULT 15

A82608

plasmid stabilization protein XF2031 [imported] - *Xylella fastidiosa* (strain 9a5c)

C;Species: *Xylella fastidiosa*

C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 03-Nov-2000

C;Accession: A82608

R;anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing and Analysis, Sao Paulo, Brazil.

Nature 406, 151-157, 2000

A;Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A;Reference number: A82515; MUID:20365717; PMID:10910347

A;Note: for a complete list of authors see reference number A59328 below

A;Accession: A82608

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-83 <SIM>

A;Cross-references: GB:AE004021; GB:AE003849; NID:g9107139; PIDN:AAF84833.1;

GSPDB:GN00128; XFSC:XF2031

A;Experimental source: strain 9a5c

R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Alves, L.M.C.; Araya, J.E.; Baia, G.S.; Baptista, C.S.; Barros, M.H.; Bonaccorsi, E.D.; Bordin, S.; Bove, J.M.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.; Colauto, N.B.; Colombo, C.; Costa, F.F.; Costa, M.C.R.; Costa-Neto, C.M.; Coutinho, L.L.; Cristofani, M.; Dias-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme, M.; Furlan, L.R.; Garnier, M.; Goldman, G.H.; Goldman, M.H.S.; Gomes, S.L.; Gruber, A.; Ho, P.L.; Hoheisel, J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigret, F.; Lambais, M.R.; Leite, L.C.C.; Lemos, E.G.M.; Lemos, M.V.F.; Lopes, S.A.; Lopes, C.R.; Machado, J.A.; Machado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.A.L.

A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; Monteiro-Vitorello, C.B.; Moon, D.H.; Nagai, M.A.; Nascimento, A.L.T.O.; Netto, L.E.S.; Nhani Jr., A.; Nobrega, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Paris, A.; Peixoto, B.R.; Pereira, G.A.G.; Pereira Jr., H.A.; Pesquero, J.B.; Quaggio, R.B.; Roberto, P.G.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki, H.E.

A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira, J.F.; Silvestri, M.L.Z.; Siqueira, W.J.; de Souza, A.A.; de Souza, A.P.; Terenzi, M.F.; Truffi, D.; Tsai, S.M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Zago, M.A.; Zatz, M.; Meidanis, J.; Setubal, J.C.

A;Reference number: A59328

A;Contents: annotation

C;Genetics:

A;Gene: XF2031

C;Superfamily: Escherichia coli parD protein

Query Match 4.8%; Score 4; DB 2; Length 83;

Best Local Similarity 100.0%; Pred. No. 7.6e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 48 EGLA 51

||||

Db 59 EGLA 62

Search completed: January 7, 2004, 10:35:01

Job time : 24 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2004, 10:34:30 ; Search time 32 Seconds  
(without alignments)  
528.963 Million cell updates/sec

Title: US-09-936-697-6  
Perfect score: 84  
Sequence: 1 QGRSGCSSQSPMRISSEN.....SPTASSQSSATNMAIHRSQP 84

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 747907 seqs, 201509753 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5083

Minimum DB seq length: 83

Maximum DB seq length: 85

Post-processing: Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Match	Length	ID	Description
No.					

1	5	6.0	83	9	US-09-864-761-35253	Sequence 35253, A
2	5	6.0	84	11	US-09-374-046A-116	Sequence 116, App
3	5	6.0	84	12	US-10-029-386-30314	Sequence 30314, A
4	5	6.0	84	12	US-10-029-386-30332	Sequence 30332, A
5	5	6.0	85	10	US-09-764-868-1203	Sequence 1203, Ap
6	5	6.0	85	12	US-10-029-386-31286	Sequence 31286, A
7	4	4.8	83	9	US-09-726-643-119	Sequence 119, App
8	4	4.8	83	9	US-09-864-761-33791	Sequence 33791, A
9	4	4.8	83	9	US-09-864-761-45444	Sequence 45444, A
10	4	4.8	83	9	US-09-864-761-48130	Sequence 48130, A
11	4	4.8	83	9	US-09-734-569-84	Sequence 84, Appl
12	4	4.8	83	9	US-09-604-287A-62	Sequence 62, Appl
13	4	4.8	83	9	US-09-800-729-178	Sequence 178, App
14	4	4.8	83	10	US-09-339-338-62	Sequence 62, Appl
15	4	4.8	83	10	US-09-731-872-369	Sequence 369, App
16	4	4.8	83	10	US-09-808-387-34	Sequence 34, Appl
17	4	4.8	83	10	US-09-764-864-1504	Sequence 1504, Ap
18	4	4.8	83	10	US-09-764-847-599	Sequence 599, App
19	4	4.8	83	10	US-09-764-877-1545	Sequence 1545, Ap
20	4	4.8	83	10	US-09-925-300-1597	Sequence 1597, Ap
21	4	4.8	83	10	US-09-843-676-9	Sequence 9, Appli
22	4	4.8	83	10	US-09-764-868-914	Sequence 914, App
23	4	4.8	83	10	US-09-766-253-9	Sequence 9, Appli
24	4	4.8	83	10	US-09-738-626-4564	Sequence 4564, Ap
25	4	4.8	83	10	US-09-796-692-2161	Sequence 2161, Ap
26	4	4.8	83	11	US-09-438-486-9	Sequence 9, Appli
27	4	4.8	83	11	US-09-977-418-30	Sequence 30, Appl
28	4	4.8	83	11	US-09-984-271-112	Sequence 112, App
29	4	4.8	83	11	US-09-764-891-3288	Sequence 3288, Ap
30	4	4.8	83	11	US-09-764-891-3452	Sequence 3452, Ap
31	4	4.8	83	11	US-09-764-891-4547	Sequence 4547, Ap
32	4	4.8	83	11	US-09-764-891-4562	Sequence 4562, Ap
33	4	4.8	83	11	US-09-764-891-4776	Sequence 4776, Ap
34	4	4.8	83	11	US-09-977-033A-30	Sequence 30, Appl
35	4	4.8	83	11	US-09-551-621-62	Sequence 62, Appl
36	4	4.8	83	12	US-09-977-751C-30	Sequence 30, Appl
37	4	4.8	83	12	US-09-876-997-369	Sequence 369, App
38	4	4.8	83	12	US-09-882-227-200	Sequence 200, App
39	4	4.8	83	12	US-10-124-805-62	Sequence 62, Appl
40	4	4.8	83	12	US-10-301-488A-44	Sequence 44, Appl
41	4	4.8	83	12	US-10-148-687-7	Sequence 7, Appli
42	4	4.8	83	12	US-10-029-386-27573	Sequence 27573, A
43	4	4.8	83	12	US-10-029-386-30244	Sequence 30244, A
44	4	4.8	83	12	US-10-029-386-30689	Sequence 30689, A
45	4	4.8	83	12	US-10-029-386-33147	Sequence 33147, A

#### ALIGNMENTS

##### RESULT 1

US-09-864-761-35253

; Sequence 35253, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES  
USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 35253  
; LENGTH: 83  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AP000352.2  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.2  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.1  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6.8  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.7

```

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 9.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 9.7
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 8.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6
; OTHER INFORMATION: EST_HUMAN HIT: AA931620.1, EVALUE 3.70e-02
; OTHER INFORMATION: SWISSPROT HIT: P09279, EVALUE 1.60e+00
US-09-864-761-35253

```

```

Query Match          6.0%; Score 5; DB 9; Length 83;
Best Local Similarity 100.0%; Pred. No. 2e+03;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      70 SQSSA 74
        |||||
Db      72 SQSSA 76

```

# RESULT 2

```

US-09-374-046A-116
; Sequence 116, Application US/09374046A
; Publication No. US20030096951A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6075-83A
; CURRENT APPLICATION NUMBER: US/09/374,046A
; CURRENT FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 116
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-374-046A-116

```

```

Query Match          6.0%; Score 5; DB 11; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      9 QSISP 13
        |||||
Db      71 QSISP 75

```

RESULT 3

US-10-029-386-30314

```
; Sequence 30314, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
USEFUL FOR GENE
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30314
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR22.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
; OTHER INFORMATION: SWISSPROT HIT: Q15459, EVALUE 2.00e-34
US-10-029-386-30314
```

```
Query Match          6.0%; Score 5; DB 12; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

```
Qy      72 SSATN 76
        |||||
Db      66 SSATN 70
```

RESULT 4

US-10-029-386-30332

```
; Sequence 30332, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
USEFUL FOR GENE
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
```

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 30332  
; LENGTH: 84  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR7.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 46  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 18  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 45  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.9  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 31  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 11  
; OTHER INFORMATION: SWISSPROT HIT: Q92830, EVALUE 2.30e-01  
US-10-029-386-30332

Query Match 6.0%; Score 5; DB 12; Length 84;  
Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 47 EEGLA 51  
|||  
Db 34 EEGLA 38

RESULT 5

US-09-764-868-1203  
; Sequence 1203, Application US/09764868  
; Patent No. US20020168711A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PTZ32  
; CURRENT APPLICATION NUMBER: US/09/764,868  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1510  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1203  
; LENGTH: 85  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-868-1203

Query Match 6.0%; Score 5; DB 10; Length 85;  
Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 80 HRSQP 84  
|||  
Db 24 HRSQP 28

RESULT 6

US-10-029-386-31286  
; Sequence 31286, Application US/10029386



; Publication No. US20030194704A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharron G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES  
 USEFUL FOR GENE  
 ; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
 ; FILE REFERENCE: AEOMICA-X-2  
 ; CURRENT APPLICATION NUMBER: US/10/029,386  
 ; CURRENT FILING DATE: 2001-12-20  
 ; NUMBER OF SEQ ID NOS: 34288  
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 31286  
 ; LENGTH: 85  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: MAP TO AC004510.1  
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49  
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.57  
 ; OTHER INFORMATION: SWISSPROT HIT: P03356, EVALUE 1.00e-10  
 US-10-029-386-31286

Query Match 6.0%; Score 5; DB 12; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ALSVA 45  
 |||||  
 Db 70 ALSVA 74

# RESULT 7

US-09-726-643-119  
 ; Sequence 119, Application US/09726643  
 ; Patent No. US20020028449A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ruben et al.  
 ; TITLE OF INVENTION: 26 Human secreted proteins  
 ; FILE REFERENCE: PZ040P1  
 ; CURRENT APPLICATION NUMBER: US/09/726,643  
 ; CURRENT FILING DATE: 2000-12-01  
 ; PRIOR APPLICATION NUMBER: PCT/US00/15187  
 ; PRIOR FILING DATE: 2000-06-02  
 ; PRIOR APPLICATION NUMBER: 60/137,725  
 ; PRIOR FILING DATE: 1999-06-07  
 ; NUMBER OF SEQ ID NOS: 190  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 119  
 ; LENGTH: 83  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-726-643-119

Query Match 4.8%; Score 4; DB 9; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 2e+04;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 SISP 13  
|||  
Db 60 SISP 63

RESULT 8

US-09-864-761-33791

; Sequence 33791, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES  
USEFUL FOR

; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

; FILE REFERENCE: Aeomica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

```

; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33791
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AJ010597.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 6.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.8
; OTHER INFORMATION: EST_HUMAN HIT: BE390947.1, EVALUE 1.00e-28
; OTHER INFORMATION: EST_HUMAN HIT: BE728616.1, EVALUE 2.00e-28
; OTHER INFORMATION: SWISSPROT HIT: O85709, EVALUE 1.60e+00
US-09-864-761-33791

```

```

Query Match          4.8%; Score 4; DB 9; Length 83;
Best Local Similarity 100.0%; Pred. No. 2e+04;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      30 QKSR 33
        ||||
Db      19 QKSR 22

```

# RESULT 9

```

US-09-864-761-45444
; Sequence 45444, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359

```

```

; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45444
;   LENGTH: 83
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL121586.28
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.52
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.68
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.72
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.69
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.76
;   OTHER INFORMATION: EST_HUMAN HIT: BE783272.1, EVALUE 5.00e-30
;   OTHER INFORMATION: SWISSPROT HIT: P24686, EVALUE 8.80e+00
US-09-864-761-45444

```

```

Query Match          4.8%; Score 4; DB 9; Length 83;
Best Local Similarity 100.0%; Pred. No. 2e+04;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      57 CLRL 60
      ||||
Db      20 CLRL 23

```

```

RESULT 10
US-09-864-761-48130
; Sequence 48130, Application US/09864761
; Patent No. US20020048763A1

```

; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES  
USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 48130  
; LENGTH: 83  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL021068.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7  
; OTHER INFORMATION: EST\_HUMAN HIT: BE972461.1, EVALUATE 4.00e-32

; OTHER INFORMATION: SWISSPROT HIT: Q09309, EVALUE 4.70e-02  
US-09-864-761-48130

Query Match 4.8%; Score 4; DB 9; Length 83;  
Best Local Similarity 100.0%; Pred. No. 2e+04;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 SENS 21  
|||  
Db 61 SENS 64

RESULT 11

US-09-734-569-84

; Sequence 84, Application US/09734569

; Patent No. US20020064816A1

; GENERAL INFORMATION:

; APPLICANT: Lerchl, Jens

; APPLICANT: Renz, Andreas

; APPLICANT: Ehrhardt, Thomas

; APPLICANT: Reindl, Andreas

; APPLICANT: Cirpus, Petra

; APPLICANT: Bischoff, Friedrich

; APPLICANT: Frank, Markus

; APPLICANT: Freund, Annette

; APPLICANT: Duwenig, Elke

; APPLICANT: Schmidt, Ralf-Michael

; APPLICANT: Reski, Ralf

; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involved

; TITLE OF INVENTION: in the synthesis of carbohydrates

; FILE REFERENCE: BASF-NAE-1332-99-US

; CURRENT APPLICATION NUMBER: US/09/734,569

; CURRENT FILING DATE: 2001-05-24

; PRIOR APPLICATION NUMBER: US 60/171,101

; PRIOR FILING DATE: 1999-12-16

; NUMBER OF SEQ ID NOS: 181

; SOFTWARE: PatentIn Ver. 2.1/WordPerfect 6.1

; SEQ ID NO 84

; LENGTH: 83

; TYPE: PRT

; ORGANISM: Physcomitrella patens

US-09-734-569-84

Query Match 4.8%; Score 4; DB 9; Length 83;  
Best Local Similarity 100.0%; Pred. No. 2e+04;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 AIHR 81  
|||  
Db 59 AIHR 62

RESULT 12

US-09-604-287A-62

; Sequence 62, Application US/09604287A

; Patent No. US20020064872A1

```
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 62
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-604-287A-62
```

```
Query Match          4.8%; Score 4; DB 9; Length 83;
Best Local Similarity 100.0%; Pred. No. 2e+04;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      15 RSIS 18
        ||||
Db      28 RSIS 31
```

# RESULT 13

```
US-09-800-729-178
; Sequence 178, Application US/09800729
; Patent No. US20020068319A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: PZ044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 178
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-178
```

```
Query Match          4.8%; Score 4; DB 9; Length 83;
Best Local Similarity 100.0%; Pred. No. 2e+04;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      49 GLAW 52
        ||||
```

Db 15 GLAW 18

RESULT 14

US-09-339-338-62

; Sequence 62, Application US/09339338A  
; Patent No. US20020102602A1  
; GENERAL INFORMATION:  
; APPLICANT: Yuqiu, Jiang  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Xu, Jiangchun  
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.470C2  
; CURRENT APPLICATION NUMBER: US/09/339,338A  
; CURRENT FILING DATE: 1999-06-23  
; NUMBER OF SEQ ID NOS: 315  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 62  
; LENGTH: 83  
; TYPE: PRT  
; ORGANISM: Homo sapien

US-09-339-338-62

Query Match 4.8%; Score 4; DB 10; Length 83;  
Best Local Similarity 100.0%; Pred. No. 2e+04;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 RSIS 18

||||

Db 28 RSIS 31

RESULT 15

US-09-731-872-369

; Sequence 369, Application US/09731872  
; Patent No. US20020102604A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED  
PROTEINS  
; FILE REFERENCE: 78.US3.REG  
; CURRENT APPLICATION NUMBER: US/09/731,872  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; PRIOR FILING DATE: 2000-03-06  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 369  
; LENGTH: 83  
; TYPE: PRT  
; ORGANISM: Homo sapiens



; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: -40...-1  
US-09-731-872-369

Query Match 4.8%; Score 4; DB 10; Length 83;  
Best Local Similarity 100.0%; Pred. No. 2e+04;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 SENS 21  
|||  
Db 75 SENS 78

Search completed: January 7, 2004, 10:39:50  
Job time : 32 secs

OM protein - protein search, using sw model

Run on: January 7, 2004, 10:31:09 ; Search time 34 Seconds  
(without alignments)  
637.542 Million cell updates/sec

Title: US-09-936-697-6  
Perfect score: 84  
Sequence: 1 QGRSGCSSQSISPMRSISEN.....SPTASSQSSATNMAIHRSQP 84

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6755

Minimum DB seq length: 83  
Maximum DB seq length: 85

Post-processing: Listing first 45 summaries

Database : SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp Vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query	
No.	Score	Match Length DB ID Description
-----		

1	6	7.1	83	5	Q8IFH5	Q8ifh5 trypanosoma
2	5	6.0	83	2	Q9AGH7	Q9agh7 klebsiella
3	5	6.0	83	9	Q8W5X5	Q8w5x5 bacterioph
4	5	6.0	83	12	Q98194	Q98194 molluscum c
5	5	6.0	83	16	Q92LP2	Q92lp2 rhizobium m
6	5	6.0	84	2	Q49621	Q49621 mycobacteri
7	5	6.0	84	5	O46279	O46279 drosophila
8	5	6.0	84	5	O46293	O46293 drosophila
9	5	6.0	84	5	O46305	O46305 drosophila
10	5	6.0	84	5	O46285	O46285 drosophila
11	5	6.0	84	5	O46299	O46299 drosophila
12	5	6.0	84	5	O46289	O46289 drosophila
13	5	6.0	84	5	O46283	O46283 drosophila
14	5	6.0	84	5	O46297	O46297 scaptomyza
15	5	6.0	84	5	O46301	O46301 drosophila
16	5	6.0	84	5	O46277	O46277 drosophila
17	5	6.0	84	5	O46295	O46295 drosophila
18	5	6.0	84	5	O46287	O46287 drosophila
19	5	6.0	84	5	O46281	O46281 drosophila
20	5	6.0	84	5	O46291	O46291 drosophila
21	5	6.0	84	5	O46275	O46275 drosophila
22	5	6.0	85	16	Q8R9D7	Q8r9d7 thermoanaer
23	4	4.8	83	1	Q8X253	Q8x253 halobacteri
24	4	4.8	83	2	Q54996	Q54996 streptomyce
25	4	4.8	83	2	Q9ZEH6	Q9zeh6 streptococc
26	4	4.8	83	2	Q8L1C4	Q8l1c4 pasteurella
27	4	4.8	83	2	O85420	O85420 pseudomonas
28	4	4.8	83	2	O85436	O85436 pseudomonas
29	4	4.8	83	2	O85432	O85432 pseudomonas
30	4	4.8	83	2	Q8L223	Q8l223 salmonella
31	4	4.8	83	2	O85443	O85443 pseudomonas
32	4	4.8	83	2	Q9S1K7	Q9s1k7 streptococc
33	4	4.8	83	2	O85442	O85442 pseudomonas
34	4	4.8	83	2	O85440	O85440 pseudomonas
35	4	4.8	83	2	O85434	O85434 pseudomonas
36	4	4.8	83	2	O85437	O85437 pseudomonas
37	4	4.8	83	2	O85416	O85416 pseudomonas
38	4	4.8	83	2	O85441	O85441 pseudomonas
39	4	4.8	83	2	O85430	O85430 pseudomonas
40	4	4.8	83	2	O85418	O85418 pseudomonas
41	4	4.8	83	2	O85433	O85433 pseudomonas
42	4	4.8	83	2	O85444	O85444 pseudomonas
43	4	4.8	83	2	Q9F4Y1	Q9f4y1 arthrospira
44	4	4.8	83	2	O85438	O85438 pseudomonas
45	4	4.8	83	2	O85412	O85412 pseudomonas

#### ALIGNMENTS

##### RESULT 1

Q8IFH5

ID Q8IFH5 PRELIMINARY; PRT; 83 AA.

AC Q8IFH5;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Hypothetical protein, unlikely.  
GN TB927.1.310.  
OS Trypanosoma brucei.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
OX NCBI\_TaxID=5691;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TREU927;  
RA Hall N., Berriman M., Lennard N.J., Harris B.R., Gerrard C.S.,  
RA Atkin R.J., Barron A.J., Bart-Delabesse E.N., Bowman S.,  
RA Bray-Allen S.P., Bringaud F., Clark L.N., Corton C.H., Cronin A.,  
RA Davies R., Doggett J., Fraser A., Gruter E., Hall S., Harper D.A.,  
RA Hertz-Fowler C., Kay M.P., Leech V., Mayes R., Price C., Quail M.A.,  
RA Rabbinoiwitsch E., Rutherford K., Sasse J., Sharp S., Shownkeen R.,  
RA Gull K., Barrell B.G., Melville S.E.;  
RT "The sequence and analysis of the highly polymorphic chromosome I of  
RT the African trypanosome, Trypanosoma brucei.";  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL929603; CAD53034.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 83 AA; 9416 MW; F8AA34BF4053BB2A CRC64;.

Query Match 7.1%; Score 6; DB 5; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 SPTASS 70  
|||||  
Db 66 SPTASS 71

## RESULT 2

### Q9AGH7

ID Q9AGH7 PRELIMINARY; PRT; 83 AA.  
AC Q9AGH7;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Hypothetical 9.5 kDa protein (Fragment).  
OS Klebsiella aerogenes.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Klebsiella.  
OX NCBI\_TaxID=28451;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21172894; PubMed=11274137;  
RA Janes B.K., Pomposiello P.J., Perez-Matos A., Najarian D.J.,  
RA Goss T.J., Bender R.A.;  
RT "Growth Inhibition Caused by Overexpression of the Structural Gene for  
RT Glutamate Dehydrogenase (gdhA) from Klebsiella aerogenes.";  
RL J. Bacteriol. 183:2709-2714(2001).  
DR EMBL; AF332586; AAK17985.1; -.  
DR InterPro; IPR000086; NUDIX\_hydrolase.  
DR Pfam; PF00293; NUDIX; 1.  
KW Hypothetical protein.  
FT NON\_TER 1 1  
SQ SEQUENCE 83 AA; 9495 MW; 1CF734940A93A529 CRC64;

Query Match 6.0%; Score 5; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 PTEAL 42  
|||||  
Db 54 PTEAL 58

RESULT 3

Q8W5X5

ID Q8W5X5 PRELIMINARY; PRT; 83 AA.  
AC Q8W5X5;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Gp32 protein.  
OS Bacteriophage PSA.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
OX NCBI\_TaxID=171618;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Loessner M.J., Sattelberger E., Zimmer M., Calendar R., Inman R.B.,  
RA Scherer S.;  
RT "Molecular analysis of *Listeria monocytogenes* ScottA bacteriophage PSA  
RT reveals ribosomal frameshifting as a general mechanism for generation  
RT of major structural proteins."  
RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AJ312240; CAC85590.1; -.  
SQ SEQUENCE 83 AA; 9589 MW; DB74D32BA7395AFD CRC64;

Query Match 6.0%; Score 5; DB 9; Length 83;  
Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 RVIEN 37  
|||||  
Db 37 RVIEN 41

RESULT 4

Q98194

ID Q98194 PRELIMINARY; PRT; 83 AA.  
AC Q98194;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE MC026L.  
GN MC026L.  
OS Molluscum contagiosum virus subtype 1 (MCVI).  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OC Molluscipoxvirus.  
OX NCBI\_TaxID=10280;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96325459; PubMed=8670425;

RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,  
 RA Moss B.;  
 RT "Genome sequence of a human tumorigenic poxvirus: Prediction of  
 RT specific host response-evasion genes."  
 RL Science 273:813-816(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,  
 RA Moss B.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; U60315; AAC55154.1; -.  
 DR InterPro; IPR007087; Znf\_C2H2.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 1.  
 SQ SEQUENCE 83 AA; 9839 MW; F27B5FEFC62FCB33 CRC64;

Query Match 6.0%; Score 5; DB 12; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 KGCLR 59  
 |||||  
 Db 32 KGCLR 36

# RESULT 5

## Q92LP2

ID Q92LP2 PRELIMINARY; PRT; 83 AA.  
 AC Q92LP2;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein R02993.  
 GN R02993 OR SMC03100.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,  
 RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 RT Sinorhizobium meliloti strain 1021."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 DR EMBL; AL591792; CAC47572.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 83 AA; 8296 MW; F249C6C83BBB6A7B CRC64;

Query Match 6.0%; Score 5; DB 16; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ALSVA 45  
|||||  
Db 11 ALSVA 15

RESULT 6

Q49621

ID Q49621 PRELIMINARY; PRT; 84 AA.  
AC Q49621;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Kas.  
OS Mycobacterium leprae.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1769;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Smith D.R.;  
RL Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Robison K.;  
RL Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; U00010; AAA17064.1; -.  
DR InterPro; IPR000794; Ketoacyl-synt.  
DR Pfam; PF00109; ketoacyl-synt; 1.  
SQ SEQUENCE 84 AA; 9237 MW; E3716511D3307DC2 CRC64;

Query Match 6.0%; Score 5; DB 2; Length 84;  
Best Local Similarity 100.0%; Pred. No. 2.2e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 SSQSS 73  
|||||  
Db 16 SSQSS 20

RESULT 7

O46279

ID O46279 PRELIMINARY; PRT; 84 AA.  
AC O46279;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Wingless (Fragment).  
GN WG OR WNT.  
OS Drosophila crassifemur (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=13053;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21971897; PubMed=11975337;  
RA Baker R.H., Desalle R.;

RL Syst. Biol. 46:654-673(1997).  
DR EMBL; U94558; AAC03310.1; -.  
DR FlyBase; FBgn0024080; Dcra\wg.  
DR InterPro; IPR005817; Wnt.  
DR Pfam; PF00110; wnt; 1.  
DR SMART; SM00097; WNT1; 1.  
FT NON\_TER 1 1  
FT NON\_TER 84 84  
SQ SEQUENCE 84 AA; 9630 MW; 40FEA9EDB48FDF1F CRC64;

Query Match 6.0%; Score 5; DB 5; Length 84;  
Best Local Similarity 100.0%; Pred. No. 2.2e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 LGTHG 64  
|||||  
Db 40 LGTHG 44

#### RESULT 8

O46293

ID O46293 PRELIMINARY; PRT; 84 AA.  
AC O46293;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Wingless (Fragment).  
GN WG OR WNT.  
OS Drosophila mulleri (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7231;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21971897; PubMed=11975337;  
RA Baker R.H., Desalle R.;  
RL Syst. Biol. 46:654-673(1997).  
DR EMBL; U94572; AAC03324.1; -.  
DR FlyBase; FBgn0023754; Dmul\wg.  
DR InterPro; IPR005817; Wnt.  
DR Pfam; PF00110; wnt; 1.  
DR SMART; SM00097; WNT1; 1.  
FT NON\_TER 1 1  
FT NON\_TER 84 84  
SQ SEQUENCE 84 AA; 9517 MW; F03A537DB493C2DE CRC64;

Query Match 6.0%; Score 5; DB 5; Length 84;  
Best Local Similarity 100.0%; Pred. No. 2.2e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 LGTHG 64  
|||||  
Db 40 LGTHG 44

#### RESULT 9



O46305

ID O46305 PRELIMINARY; PRT; 84 AA.  
AC O46305;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Wingless (Fragment).  
GN WG OR WNT.  
OS Drosophila dasyncnemia (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=58308;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21971897; PubMed=11975337;  
RA Baker R.H., Desalle R.;  
RL Syst. Biol. 46:654-673(1997).  
DR EMBL; U94584; AAC03336.1; -.  
DR FlyBase; FBgn0024063; Ddas\wg.  
DR InterPro; IPR005817; Wnt.  
DR Pfam; PF00110; wnt; 1.  
DR SMART; SM00097; WNT1; 1.  
FT NON\_TER 1 1  
FT NON\_TER 84 84  
SQ SEQUENCE 84 AA; 9598 MW; F4FEBC2C8523C2DE CRC64;

Query Match 6.0%; Score 5; DB 5; Length 84;  
Best Local Similarity 100.0%; Pred. No. 2.2e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 LGTHG 64  
|||||  
Db 40 LGTHG 44

RESULT 10

O46285

ID O46285 PRELIMINARY; PRT; 84 AA.  
AC O46285;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Wingless (Fragment).  
GN WG OR WNT.  
OS Drosophila iki (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=58311;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21971897; PubMed=11975337;  
RA Baker R.H., Desalle R.;  
RL Syst. Biol. 46:654-673(1997).  
DR EMBL; U94564; AAC03316.1; -.  
DR FlyBase; FBgn0023823; Diki\wg.

DR InterPro; IPR005817; Wnt.  
DR Pfam; PF00110; wnt; 1.  
DR SMART; SM00097; WNT1; 1.  
FT NON\_TER 1 1  
FT NON\_TER 84 84  
SQ SEQUENCE 84 AA; 9476 MW; E1FEA93C60EC486E CRC64;

Query Match 6.0%; Score 5; DB 5; Length 84;  
Best Local Similarity 100.0%; Pred. No. 2.2e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 LGTHG 64  
|||||  
Db 40 LGTHG 44

#### RESULT 11

O46299

ID O46299 PRELIMINARY; PRT; 84 AA.  
AC O46299;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Wingless (Fragment).  
GN WG OR WNT.  
OS Drosophila silvestris (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=47010;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21971897; PubMed=11975337;  
RA Baker R.H., Desalle R.;  
RL Syst. Biol. 46:654-673(1997).  
DR EMBL; U94578; AAC03330.1; -.  
DR FlyBase; FBgn0023619; Dsil\wg.  
DR InterPro; IPR005817; Wnt.  
DR Pfam; PF00110; wnt; 1.  
FT NON\_TER 1 1  
FT NON\_TER 84 84  
SQ SEQUENCE 84 AA; 9579 MW; F4FEBC2C95F9A2DE CRC64;

Query Match 6.0%; Score 5; DB 5; Length 84;  
Best Local Similarity 100.0%; Pred. No. 2.2e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 LGTHG 64  
|||||  
Db 40 LGTHG 44

#### RESULT 12

O46289

ID O46289 PRELIMINARY; PRT; 84 AA.  
AC O46289;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)

DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Wingless (Fragment).  
 GN WG OR WNT.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21971897; PubMed=11975337;  
 RA Baker R.H., Desalle R.;  
 RL Syst. Biol. 46:654-673(1997).  
 DR EMBL; U94568; AAC03320.1; -.  
 DR FlyBase; FBgn0023554; Dwts\wg.  
 DR InterPro; IPR005817; Wnt.  
 DR Pfam; PF00110; wnt; 1.  
 DR SMART; SM00097; WNT1; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 84 84  
 SQ SEQUENCE 84 AA; 9494 MW; F4FEB862C54DF78F CRC64;  
  
 Query Match 6.0%; Score 5; DB 5; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 LGTHG 64  
 |||||  
 Db 40 LGTHG 44

RESULT 13  
 O46283

ID O46283 PRELIMINARY; PRT; 84 AA.  
 AC O46283;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Wingless (Fragment).  
 GN WG OR WNT.  
 OS Drosophila disjuncta (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=46796;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21971897; PubMed=11975337;  
 RA Baker R.H., Desalle R.;  
 RL Syst. Biol. 46:654-673(1997).  
 DR EMBL; U94562; AAC03314.1; -.  
 DR FlyBase; FBgn0024055; Ddis\wg.  
 DR InterPro; IPR005817; Wnt.  
 DR Pfam; PF00110; wnt; 1.  
 DR SMART; SM00097; WNT1; 1.  
 FT NON\_TER 1 1

FT NON\_TER 84 84  
SQ SEQUENCE 84 AA; 9511 MW; EFE41857B94315DE CRC64;

Query Match 6.0%; Score 5; DB 5; Length 84;  
Best Local Similarity 100.0%; Pred. No. 2.2e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 LGTHG 64  
|||||  
Db 40 LGTHG 44

RESULT 14

O46297

ID O46297 PRELIMINARY; PRT; 84 AA.  
AC O46297;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Wingless (Fragment).  
GN WG OR WNT.  
OS Scaptomyza albobittata.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Scaptomyza.  
OX NCBI\_TaxID=7299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21971897; PubMed=11975337;  
RA Baker R.H., Desalle R.;  
RL Syst. Biol. 46:654-673(1997).  
DR EMBL; U94576; AAC03328.1; -.  
DR FlyBase; FBgn0023481; Salb\wg.  
DR InterPro; IPR005817; Wnt.  
DR Pfam; PF00110; wnt; 1.  
FT NON\_TER 1 1  
FT NON\_TER 84 84  
SQ SEQUENCE 84 AA; 9547 MW; F4FEA9EDB493C2DE CRC64;

Query Match 6.0%; Score 5; DB 5; Length 84;  
Best Local Similarity 100.0%; Pred. No. 2.2e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 LGTHG 64  
|||||  
Db 40 LGTHG 44

RESULT 15

O46301

ID O46301 PRELIMINARY; PRT; 84 AA.  
AC O46301;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Wingless (Fragment).  
GN WG OR WNT.

OS *Drosophila soonae* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=58310;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21971897; PubMed=11975337;  
 RA Baker R.H., Desalle R.;  
 RL Syst. Biol. 46:654-673(1997).  
 DR EMBL; U94580; AAC03332.1; -.  
 DR FlyBase; FBgn0023609; Dsoo\wg.  
 DR InterPro; IPR005817; Wnt.  
 DR Pfam; PF00110; wnt; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 84 84  
 SQ SEQUENCE 84 AA; 9547 MW; F4FEA9EDB493C2DE CRC64;

Query Match 6.0%; Score 5; DB 5; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 LGTHG 64  
 |||||  
 Db 40 LGTHG 44

Search completed: January 7, 2004, 10:34:27  
 Job time : 36 secs

OM protein - protein search, using sw model

Run on: January 7, 2004, 10:28:14 ; Search time 11 Seconds  
(without alignments)  
359.113 Million cell updates/sec

Title: US-09-936-697-6  
Perfect score: 84  
Sequence: 1 QGRSGCSSQSISPMRSISEN.....SPTASSQSSATNMAIHRSP 84

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 592

Minimum DB seq length: 83

Maximum DB seq length: 85

Post-processing: Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	5	6.0	85	1	GLR1_ECOLI	P00277 escherichia
2	5	6.0	85	1	PYS1_FREDI	P11397 fremyella d
3	4	4.8	83	1	ELAC_MACEU	O62845 macropus eu
4	4	4.8	83	1	EX7S_RHIME	Q92ri9 rhizobium m
5	4	4.8	83	1	IBB3_SOYBN	P01064 glycine max
6	4	4.8	83	1	MULI_PSEAE	P11221 pseudomonas
7	4	4.8	83	1	PSK4_ORYSA	Q9ar88 oryza sativ
8	4	4.8	83	1	RR17_PORPU	P51305 porphyra pu
9	4	4.8	83	1	TMOB_PSEME	Q00457 pseudomonas
10	4	4.8	84	1	COAB_BPI22	P15416 bacterioph
11	4	4.8	84	1	CPB2_ECOLI	P03847 escherichia
12	4	4.8	84	1	DEFI_TENMO	Q27023 tenebrio mo
13	4	4.8	84	1	EX7S_CAUCR	Q9a6m3 caulobacter
14	4	4.8	84	1	EX7S_HAEIN	P43914 haemophilus
15	4	4.8	84	1	GVM1_HALN1	P24377 halobacteri
16	4	4.8	84	1	IAC2_HUMAN	P20155 homo sapien
17	4	4.8	84	1	IM13_ORYSA	Q9xgy5 oryza sativ

18	4	4.8	84	1	MT22_ORYSA	P93433	oryza sativ
19	4	4.8	84	1	RL23_HALMA	P12732	haloarcula
20	4	4.8	84	1	RS16_RALSO	Q8y0w0	ralstonia s
21	4	4.8	84	1	TOLB_ACTPL	Q44156	actinobacil
22	4	4.8	84	1	WHIP_STRAW	Q93jl8	streptomyce
23	4	4.8	84	1	Y2F4_STRCO	Q9xa10	streptomyce
24	4	4.8	84	1	YFAE_ECOLI	P37910	escherichia
25	4	4.8	85	1	ACTO_ENTHI	O15602	entamoeba h
26	4	4.8	85	1	EL2A_HORSE	P37357	equus cabal
27	4	4.8	85	1	FDCS_HUMAN	Q8nfu4	homo sapien
28	4	4.8	85	1	HEPC_MORCS	P82951	morone chry
29	4	4.8	85	1	HKL5_MAIZE	P56663	zea mays (m
30	4	4.8	85	1	HLG3_MAIZE	P56669	zea mays (m
31	4	4.8	85	1	NEU1_PAPHA	P32005	papio hamad
32	4	4.8	85	1	PTHP_BUCBP	Q89b03	buchnera ap
33	4	4.8	85	1	RR12_SOLNI	Q9xqe1	solanum nig
34	4	4.8	85	1	RS17_MYCCA	P10131	mycoplasma
35	4	4.8	85	1	YOR4_BPSPP	Q38440	bacteriopha
36	3	3.6	83	1	ACP_ANASP	P58553	anabaena sp
37	3	3.6	83	1	APC1_HUMAN	P02654	homo sapien
38	3	3.6	83	1	APC1_PAPHA	P34929	papio hamad
39	3	3.6	83	1	ASR_SALTI	Q93mh4	salmonella
40	3	3.6	83	1	ATP9_HELAN	P17254	helianthus
41	3	3.6	83	1	ATPH_GALSU	P35013	galdieria s
42	3	3.6	83	1	ATPH_PAVLU	P28530	pavlova lut
43	3	3.6	83	1	BLE2_STAAU	P22491	staphylococ
44	3	3.6	83	1	C554_PARSP	P00105	paracoccus
45	3	3.6	83	1	CALD_BOVIN	Q27976	bos taurus

# ALIGNMENTS

## RESULT 1

### GLR1\_ECOLI

ID GLR1\_ECOLI STANDARD; PRT; 85 AA.  
AC P00277;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Glutaredoxin 1 (Grx1).  
GN GRXA OR GRX OR B0849 OR SF0802.  
OS Escherichia coli, and  
OS Shigella flexneri.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562, 623;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli;  
RX MEDLINE=87005940; PubMed=3530878;  
RA Hoeoeg J.-O., von Bahr-Lindstrom H., Joernvall H., Holmgren A.;  
RT "Cloning and expression of the glutaredoxin (grx) gene of Escherichia  
RT coli.";  
RL Gene 43:13-21(1986).  
RN [2]  
RP SEQUENCE FROM N.A.

RC SPECIES=E.coli;  
 RA Chatterjee P.K., Sternberg N.L.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=E.coli; STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=E.coli; STRAIN=K12;  
 RX MEDLINE=97061202; PubMed=8905232;  
 RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
 RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,  
 RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
 RA Yano M., Horiuchi T.;  
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome  
 RT corresponding to the 12.7-28.0 min region on the linkage map.";  
 RL DNA Res. 3:137-155(1996).  
 RN [5]  
 RP SEQUENCE.  
 RC SPECIES=E.coli; STRAIN=K12;  
 RX MEDLINE=84004402; PubMed=6352262;  
 RA Hoeoeg J.-O., Joernvall H., Holmgren A., Carlquist M., Persson M.;  
 RT "The primary structure of Escherichia coli glutaredoxin. Distant  
 RT homology with thioredoxins in a superfamily of small proteins with a  
 RT redox-active cystine disulfide/cysteine dithiol.";  
 RL Eur. J. Biochem. 136:223-232(1983).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;  
 RX MEDLINE=22272406; PubMed=12384590;  
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,  
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
 RA Yu J.;  
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity  
 RT through comparison with genomes of Escherichia coli K12 and O157.";  
 RL Nucleic Acids Res. 30:4432-4441(2002).  
 RN [7]  
 RP STRUCTURE BY NMR.  
 RC SPECIES=E.coli;  
 RX MEDLINE=91364685; PubMed=1889405;  
 RA Sodano P., Chary K.V.R., Bjoernberg O., Holmgren A., Kren B.,  
 RA Fuchs J.A., Wuethrich K.;  
 RT "Nuclear magnetic resonance studies of recombinant Escherichia coli  
 RT glutaredoxin. Sequence-specific assignments and secondary structure  
 RT determination of the oxidized form.";  
 RL Eur. J. Biochem. 200:369-377(1991).



RN [8]  
 RP STRUCTURE BY NMR.  
 RC SPECIES=E.coli;  
 RX MEDLINE=92046066; PubMed=1942053;  
 RA Sodano P., Xia T.-H., Bushweller J.H., Bjoernberg O., Holmgren A.,  
 RA Billeter M., Wuethrich K.;  
 RT "Sequence-specific 1H NMR assignments and determination of the three-  
 RT dimensional structure of reduced Escherichia coli glutaredoxin.";  
 RL J. Mol. Biol. 221:1311-1324(1991).  
 RN [9]  
 RP STRUCTURE BY NMR.  
 RC SPECIES=E.coli;  
 RX MEDLINE=93278264; PubMed=1304339;  
 RA Xia T.-H., Bushweller J.H., Sodano P., Billeter M., Bjoernberg O.,  
 RA Holmgren A., Wuethrich K.;  
 RT "NMR structure of oxidized Escherichia coli glutaredoxin: comparison  
 RT with reduced E. coli glutaredoxin and functionally related  
 RT proteins.";  
 RL Protein Sci. 1:310-321(1992).  
 RN [10]  
 RP STRUCTURE BY NMR.  
 RC SPECIES=E.coli;  
 RX MEDLINE=97270442; PubMed=9125525;  
 RA Kelley J.J. III, Caputo M., Eaton S.F., Laue T.M., Bushweller J.H.;  
 RT "Comparison of backbone dynamics of reduced and oxidized Escherichia  
 RT coli glutaredoxin-1 using 15N NMR relaxation measurements.";  
 RL Biochemistry 36:5029-5044(1997).  
 CC -!- FUNCTION: THE DISULFIDE BOND FUNCTIONS AS AN ELECTRON CARRIER IN  
 CC THE GLUTATHIONE-DEPENDENT SYNTHESIS OF DEOXYRIBONUCLEOTIDES BY THE  
 CC ENZYME RIBONUCLEOTIDE REDUCTASE. IN ADDITION, IT IS ALSO INVOLVED  
 CC IN REDUCING SOME DISULFIDES IN A COUPLED SYSTEM WITH GLUTATHIONE  
 CC REDUCTASE.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- SIMILARITY: BELONGS TO THE GLUTAREDOXIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M13449; AAA23936.1; -.  
 DR EMBL; U18655; AAC43449.1; -.  
 DR EMBL; AE000187; AAC73936.1; -.  
 DR EMBL; D90722; BAA35552.1; -.  
 DR EMBL; D90723; BAA35560.1; -.  
 DR EMBL; AE015109; AAN42435.1; ALT\_INIT.  
 DR PIR; A00283; GDEC.  
 DR PDB; 1EGO; 31-OCT-93.  
 DR PDB; 1EGR; 31-OCT-93.  
 DR PDB; 1GRX; 24-JUN-98.  
 DR PDB; 1QFN; 01-JAN-00.  
 DR ECO2DBASE; B011.0; 6TH EDITION.  
 DR EcoGene; EG10417; grxA.  
 DR InterPro; IPR002109; Glutaredoxin.

DR Pfam; PF00462; glutaredoxin; 1.  
 DR PRINTS; PR00160; GLUTAREDOXIN.  
 DR PROSITE; PS00195; GLUTAREDOXIN; 1.  
 KW Redox-active center; Electron transport; 3D-structure;  
 KW Deoxyribonucleotide synthesis; Complete proteome.  
 FT DISULFID 11 14 REDOX-ACTIVE.  
 FT STRAND 2 6  
 FT HELIX 12 27  
 FT STRAND 32 36  
 FT HELIX 38 41  
 FT TURN 42 42  
 FT HELIX 45 52  
 FT TURN 53 53  
 FT STRAND 61 64  
 FT TURN 65 66  
 FT STRAND 67 70  
 FT HELIX 72 82  
 FT TURN 83 85  
 SQ SEQUENCE 85 AA; 9685 MW; 33C185A47021EF42 CRC64;

Query Match 6.0%; Score 5; DB 1; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRSGC 6  
 |||||  
 Db 7 GRSGC 11

## RESULT 2

### PYS1\_FREDI

ID PYS1\_FREDI STANDARD; PRT; 85 AA.  
 AC P11397;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE Phycobilisome 9.7 kDa linker polypeptide, phycocyanin-associated, rod  
 DE (L-9.7/R) (Rod capping linker protein).  
 GN CPCD2.  
 OS Fremyella diplosiphon (Calothrix PCC 7601).  
 OC Bacteria; Cyanobacteria; Nostocales; Microchaetaceae; Fremyella.  
 OX NCBI\_TaxID=1197;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87222193; PubMed=3108238;  
 RA Lomax T.L., Conley P.B., Schilling J., Grossman A.R.;  
 RT "Isolation and characterization of light-regulated phycobilisome  
 RT linker polypeptide genes and their transcription as a polycistronic  
 RT mRNA."  
 RL J. Bacteriol. 169:2675-2684(1987).  
 CC -!- SIMILARITY: TO OTHER PHYCOBILISOME LINKER PROTEINS.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M16490; AAA24888.1; -.  
 DR InterPro; IPR001685; CpcD-like.  
 DR Pfam; PF01383; CpcD; 1.  
 DR ProDom; PD002828; CpcD-like\_C; 1.  
 KW Phycobilisome; Photosynthesis.  
 SQ SEQUENCE 85 AA; 9745 MW; 65FD1CF6A4964AE8 CRC64;

Query Match 6.0%; Score 5; DB 1; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 ENPTE 40  
 |||||  
 Db 76 ENPTE 80

### RESULT 3

#### ELAC\_MACEU

ID ELAC\_MACEU STANDARD; PRT; 83 AA.  
 AC O62845;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Early lactation protein.  
 GN ELP.  
 OS Macropus eugenii (Tammar wallaby).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.  
 OX NCBI\_TaxID=9315;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-30 AND 47-71.  
 RC TISSUE=Lactating mammary gland;  
 RX MEDLINE=99004081; PubMed=9787813;  
 RA Simpson K.J., Shaw D., Nicholas K.R.;  
 RT "Developmentally-regulated expression of a putative protease inhibitor  
 RT gene in the lactating mammary gland of the tammar wallaby, Macropus  
 RT eugenii. eugenii.";  
 RL Comp. Biochem. Physiol. 120B:535-541(1998).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: FOUND IN THE WHEY FRACTION OF MILK.  
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED ONLY IN THE EARLY STAGE OF  
 CC LACTATION. DECREASE IN EXPRESSION CORRELATES WITH A CHANGE IN THE  
 CC SUCKING PATTERN OF THE YOUNG.  
 CC -!- PTM: N-GLYCOSYLATED (PROBABLE).  
 CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AJ000490; CAA04128.1; -.  
 DR HSSP; P02760; 1BIK.  
 DR InterPro; IPR002223; Kunitz\_BPTI.  
 DR Pfam; PF00014; Kunitz\_BPTI; 1.  
 DR PRINTS; PR00759; BASICPTASE.  
 DR ProDom; PD000222; Kunitz\_BPTI; 1.  
 DR SMART; SM00131; KU; 1.  
 DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
 DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.  
 KW Serine protease inhibitor; Glycoprotein; Lactation.  
 FT DOMAIN 23 73 BPTI/KUNITZ INHIBITOR.  
 FT DISULFID 23 73 BY SIMILARITY.  
 FT DISULFID 32 56 BY SIMILARITY.  
 FT DISULFID 48 69 BY SIMILARITY.  
 FT ACT\_SITE 33 34 REACTIVE BOND (BY SIMILARITY).  
 FT DOMAIN 76 83 POLY-ASN.  
 FT CARBOHYD 14 14 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 31 31 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 83 AA; 9583 MW; CD94CD35EF3175E1 CRC64;

Query Match 4.8%; Score 4; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CSSQ 9  
 ||||  
 Db 32 CSSQ 35

#### RESULT 4

##### EX7S\_RHIME

ID EX7S\_RHIME STANDARD; PRT; 83 AA.  
 AC Q92RI9;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Probable exodeoxyribonuclease VII small subunit (EC 3.1.11.6)  
 DE (Exonuclease VII small subunit).  
 GN XSEB OR R00882 OR SMC00970.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,  
 RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 RT Sinorhizobium meliloti strain 1021."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large

CC acid-insoluble oligonucleotides, which are then degraded further  
 CC into small acid-soluble oligonucleotides (By similarity).  
 CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-  
 CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.  
 CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE XSEB FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AL591785; CAC45454.1; -.  
 DR HAMAP; MF\_00337; -; 1.  
 DR InterPro; IPR003761; Exonuc\_VII\_S.  
 DR Pfam; PF02609; Exonuc\_VII\_S; 1.  
 DR ProDom; PD028235; Exonuc\_VII\_S; 1.  
 DR TIGRFAMs; TIGR01280; xseB; 1.  
 KW Hydrolase; Nuclease; Exonuclease; Complete proteome.  
 SQ SEQUENCE 83 AA; 9259 MW; EC7FBE654F0871BC CRC64;

Query Match 4.8%; Score 4; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 AVEE 48  
 ||||  
 Db 17 AVEE 20

# RESULT 5

## IBB3\_SOYBN

ID IBB3\_SOYBN STANDARD; PRT; 83 AA.  
 AC P01064;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Bowman-Birk type proteinase inhibitor D-II precursor (IV).  
 OS Glycine max (Soybean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 OX NCBI\_TaxID=3847;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Joudrier P.E., Foard D.E., Floener L.A., Larkins B.A.;  
 RT "Isolation and sequence of cDNA encoding the soybean protease  
 RT inhibitors PI IV and C-II.";  
 RL Plant Mol. Biol. 10:35-42(1987).  
 RN [2]  
 RP SEQUENCE OF 9-83.  
 RX MEDLINE=78150870; PubMed=641033;

RA Odani S., Ikenaka T.;  
 RT "Studies on soybean trypsin inhibitors, XII. Linear sequences of two  
 RT soybean double-headed trypsin inhibitors, D-II and E-I.";  
 RL J. Biochem. 83:737-745(1978).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
 RX MEDLINE=92112932; PubMed=1730730;  
 RA Chen P., Rose J., Love R., Wei C.H., Wang B.C.;  
 RT "Reactive sites of an anticarcinogenic Bowman-Birk proteinase  
 RT inhibitor are similar to other trypsin inhibitors.";  
 RL J. Biol. Chem. 267:1990-1994(1992).  
 CC -!- SIMILARITY: BELONGS TO THE BOWMAN-BIRK SERINE PROTEASE INHIBITORS  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M20733; AAA33954.1; -.  
 DR PIR; S07941; S07941.  
 DR PDB; 1PI2; 15-APR-92.  
 DR InterPro; IPR000877; Bowman-Birk\_leg.  
 DR Pfam; PF00228; Bowman-Birk\_leg; 2.  
 DR ProDom; PD002168; Bowman-Birk\_leg; 1.  
 DR SMART; SM00269; BowB; 1.  
 DR PROSITE; PS00281; BOWMAN\_BIRK; 1.  
 KW Serine protease inhibitor; 3D-structure.  
 FT PROPEP 1 8  
 FT CHAIN 9 83 BOWMAN-BIRK TYPE PROTEINASE INHIBITOR D-  
 FT II.  
 FT ACT\_SITE 32 33 INTERACTION WITH TRYPSIN.  
 FT ACT\_SITE 59 60 INTERACTION WITH TRYPSIN.  
 FT DISULFID 24 78  
 FT DISULFID 25 40  
 FT DISULFID 28 74  
 FT DISULFID 30 38  
 FT DISULFID 48 55  
 FT DISULFID 52 67  
 FT DISULFID 57 65  
 SQ SEQUENCE 83 AA; 9468 MW; 55A5F9524373C20B CRC64;

Query Match 4.8%; Score 4; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 RSQP 84  
 ||||  
 Db 59 RSQP 62

RESULT 6  
 MULI\_PSEAE  
 ID MULI\_PSEAE STANDARD; PRT; 83 AA.

AC P11221;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Major outer membrane lipoprotein precursor (Murein-lipoprotein)  
 DE (Lipoprotein I).  
 GN OPRI OR PA2853.  
 OS *Pseudomonas aeruginosa*.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; *Pseudomonas*.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Isolate PA2;  
 RX MEDLINE=89313294; PubMed=2473376;  
 RA Cornelis P., Bouia A., Belarbi A., Guyonvarch A., Kammerer B.,  
 RA Hannaert V., Hubert J.-C.;  
 RT "Cloning and analysis of the gene for the major outer membrane  
 RT lipoprotein from *Pseudomonas aeruginosa*."  
 RL Mol. Microbiol. 3:421-428(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89327122; PubMed=2502533;  
 RA Duchene M., Barron C., Schweizer A., von Sprecht B.-U., Domdey H.;  
 RT "Pseudomonas aeruginosa outer membrane lipoprotein I gene: molecular  
 RT cloning, sequence, and expression in *Escherichia coli*."  
 RL J. Bacteriol. 171:4130-4137(1989).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PAO1;  
 RX MEDLINE=92268853; PubMed=1588307;  
 RA Saint-Onge A., Romeyer F., Lebel P., Masson L., Brousseau R.;  
 RT "Specificity of the *Pseudomonas aeruginosa* PAO1 lipoprotein I gene as  
 RT a DNA probe and PCR target region within the Pseudomonadaceae."  
 RL J. Gen. Microbiol. 138:733-741(1992).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PAO1;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an  
 RT opportunistic pathogen."  
 RL Nature 406:959-964(2000).

CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid  
 CC anchor.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X13748; CAA32013.1; -.  
 DR EMBL; M25761; AAA25880.1; -.  
 DR EMBL; X58714; CAA41550.1; -.  
 DR EMBL; A07695; CAA00707.1; -.  
 DR EMBL; AE004712; AAG06241.1; -.  
 DR PIR; A33854; A33854.  
 DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; 1.  
 KW Outer membrane; Lipoprotein; Signal; Complete proteome.  
 FT SIGNAL 1 19  
 FT CHAIN 20 83 MAJOR OUTER MEMBRANE LIPOPROTEIN.  
 FT LIPID 20 20 N-ACYL DIGLYCERIDE.  
 SQ SEQUENCE 83 AA; 8835 MW; E87F52B86B04DBA4 CRC64;

Query Match 4.8%; Score 4; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GCSS 8  
 ||||  
 Db 19 GCSS 22

# RESULT 7

## PSK4\_ORYSA

ID PSK4\_ORYSA STANDARD; PRT; 83 AA.  
 AC Q9AR88;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Phytosulfokines 4 precursor [Contains: Phytosulfokine-alpha (PSK-alpha) (Phytosulfokine-a); Phytosulfokine-beta (PSK-beta) (Phytosulfokine-b)].  
 GN PSK4.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lorbiecke R., Sauter M.;  
 RT "Precursor homologs of the PSK-alpha peptide growth factor are conserved in structure but not primary sequence: identification and characterization of the rice gene family."  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF PSK-ALPHA AND PSK-BETA, CHARACTERIZATION, AND SULFATION.  
 RX MEDLINE=20200296; PubMed=10734215;  
 RA Matsubayashi Y., Takagi L., Sakagami Y.;  
 RT "Phytosulfokine-alpha, a sulfated pentapeptide, stimulates the proliferation of rice cells by means of specific high- and low-affinity binding sites."  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:13357-13362(1997).  
 CC -!- FUNCTION: PROMOTES PLANT CELL DIFFERENTIATION, ORGANOGENESIS AND  
 CC SOMATIC EMBRYOGENESIS AS WELL AS CELL PROLIFERATION.



CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).  
 CC -!- PTM: SULFATION IS IMPORTANT FOR ACTIVITY AND FOR THE BINDING TO A  
 CC PUTATIVE MEMBRANE RECEPTOR.  
 CC -!- PTM: PSK-ALPHA IS PRODUCED BY ENDOPEPTIDASE DIGESTION. PSK-BETA IS  
 CC PRODUCED FROM PSK-ALPHA BY EXOPEPTIDASE DIGESTION.  
 CC -!- SIMILARITY: BELONGS TO THE PHYTOSULFOKINE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ276693; CAC34733.1; -.  
 DR Gramene; Q9AR88; -.  
 KW Growth factor; Differentiation; Signal; Sulfation; Multigene family.  
 FT SIGNAL 1 28 POTENTIAL.  
 FT PROPEP 29 74 POTENTIAL.  
 FT PEPTIDE 75 79 PHYTOSULFOKINE-ALPHA.  
 FT PEPTIDE 75 78 PHYTOSULFOKINE-BETA.  
 FT PROPEP 80 83 POTENTIAL.  
 FT MOD\_RES 75 75 SULFATION.  
 FT MOD\_RES 77 77 SULFATION.  
 SQ SEQUENCE 83 AA; 8574 MW; 5F7F03BF07CE6AD0 CRC64;

Query Match 4.8%; Score 4; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 72 SSAT 75  
 ||||  
 Db 21 SSAT 24

# RESULT 8

## RR17\_PORPU

ID RR17\_PORPU STANDARD; PRT; 83 AA.  
 AC P51305;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Chloroplast 30S ribosomal protein S17.  
 GN RPS17.  
 OS Porphyra purpurea.  
 OG Chloroplast.  
 OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.  
 OX NCBI\_TaxID=2787;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Avonport;  
 RA Reith M.E., Munholland J.;  
 RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast  
 RT genome."  
 RL Plant Mol. Biol. Rep. 13:333-335(1995).  
 CC -!- SIMILARITY: BELONGS TO THE S17P FAMILY OF RIBOSOMAL PROTEINS.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U38804; AAC08191.1; -.  
 DR PIR; S73226; S73226.  
 DR HSSP; P23828; 1RIP.  
 DR InterPro; IPR000266; Ribosomal\_S17.  
 DR Pfam; PF00366; Ribosomal\_S17; 1.  
 DR PRINTS; PR00973; RIBOSOMALS17.  
 DR ProDom; PD001295; Ribosomal\_S17; 1.  
 DR PROSITE; PS00056; RIBOSOMAL\_S17; 1.  
 KW Ribosomal protein; Chloroplast.  
 SQ SEQUENCE 83 AA; 9552 MW; 503A3F59802B4580 CRC64;

Query Match 4.8%; Score 4; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 VAVE 47  
 ||||  
 Db 21 VAVE 24

# RESULT 9

## TMOB\_PSEME

ID TMOB\_PSEME STANDARD; PRT; 83 AA.  
 AC Q00457;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Toluene-4-monooxygenase system protein B (EC 1.14.13.-).  
 GN TMOB.  
 OS Pseudomonas mendocina.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=300;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-27.  
 RC STRAIN=KR1;  
 RX MEDLINE=91358306; PubMed=1885512;  
 RA Yen K.-M., Karl M.R., Blatt L.M., Simon M.J., Winter R.B.,  
 RA Fausset P.R., Lu H.S., Harcourt A.A., Chen K.K.;  
 RT "Cloning and characterization of a Pseudomonas mendocina KR1 gene  
 RT cluster encoding toluene-4-monooxygenase.";  
 RL J. Bacteriol. 173:5315-5327(1991).  
 CC -!- FUNCTION: HYDROXYLATES TOLUENE TO FORM P-CRESOL.  
 CC -!- COFACTOR: FAD; requires Fe(2+) for activity.  
 CC -!- PATHWAY: Toluene degradation; first step.  
 CC -!- SUBUNIT: THE MULTICOMPONENT ENZYME TOLUENE-4-MONOOXYGENASE  
 CC IS FORMED BY THE TMOA, TMOB, TMOC, TMOD, TMOE AND TMOF  
 CC POLYPEPTIDES.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M65106; AAA26000.1; -.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Flavoprotein;
KW Monooxygenase; FAD; Iron.
FT INIT_MET 0 0
SQ SEQUENCE 83 AA; 9457 MW; 4729FEF73F266F44 CRC64;

```

```

Query Match          4.8%; Score 4; DB 1; Length 83;
Best Local Similarity 100.0%; Pred. No. 3.1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      37 NPTE 40
        ||||
Db      72 NPTE 75

```

#### RESULT 10

##### COAB\_BPI22

```

ID COAB_BPI22 STANDARD; PRT; 84 AA.
AC P15416;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Coat protein B precursor (Major coat protein).
GN VIII.
OS Bacteriophage I2-2.
OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
OX NCBI_TaxID=10869;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92211729; PubMed=1556749;
RA Stassen A.P., Schonmakers E.F., Yu M., Schoenmakers J.G.,
RA Konings R.N.H.;
RT "Nucleotide sequence of the genome of the filamentous bacteriophage
RT I2-2: module evolution of the filamentous phage genome.";
RL J. Mol. Evol. 34:141-152(1992).
CC -!- FUNCTION: COAT PROTEIN B IS THE MAJOR COAT PROTEIN OF THE VIRION.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14336; CAA32517.1; -.
DR PIR; S08090; S08090.
DR HSSP; P03620; 1IFL.

```

DR Pfam; PF05371; Phage\_Coat\_Gp8; 1.  
 KW Coat protein; Signal.  
 FT SIGNAL 1 29  
 FT CHAIN 30 84 COAT PROTEIN B.  
 SQ SEQUENCE 84 AA; 8551 MW; 409884A4E186A2EC CRC64;

Query Match 4.8%; Score 4; DB 1; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 SVAV 46  
 ||||  
 Db 65 SVAV 68

# RESULT 11

## CPB2\_ECOLI

ID CPB2\_ECOLI STANDARD; PRT; 84 AA.  
 AC P03847; P13947;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Replication regulatory protein repA2 (copB protein).  
 GN REPA2 OR COPB.  
 OS Escherichia coli.  
 OG Plasmid IncFII R100, Plasmid R6-5, Plasmid IncFII NR1, and  
 OG Plasmid IncFVI pSU212.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC PLASMID=IncFII R100;  
 RX MEDLINE=81074309; PubMed=7003300;  
 RA Rosen J., Ryder T., Inokuchi H., Ohtsubo H., Ohtsubo E.;  
 RT "Genes and sites involved in replication and incompatibility of an  
 RT R100 plasmid derivative based on nucleotide sequence analysis.";  
 RL Mol. Gen. Genet. 179:527-537(1980).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC PLASMID=IncFII R100;  
 RX MEDLINE=86319522; PubMed=3019092;  
 RA Ohtsubo H., Ryder T.B., Maeda Y., Armstrong K., Ohtsubo E.;  
 RT "DNA replication of the resistance plasmid R100 and its control.";  
 RL Adv. Biophys. 21:115-133(1986).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC PLASMID=R6-5;  
 RX MEDLINE=82060121; PubMed=7029525;  
 RA Danbara H., Brady G., Timmis J.K., Timmis K.N.;  
 RT "Regulation of DNA replication: 'target' determinant of the  
 RT replication control elements of plasmid R6-5 lies within a control  
 RT element gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 78:4699-4703(1981).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC PLASMID=IncFII NR1;

RX MEDLINE=88058738; PubMed=2445727;  
 RA Dong X., Womble D.D., Rownd R.H.;  
 RT "Transcriptional pausing in a region important for plasmid NR1  
 RT replication control.";  
 RL J. Bacteriol. 169:5353-5363(1987).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC PLASMID=IncFII NR1;  
 RX MEDLINE=85160860; PubMed=2580099;  
 RA Womble D.D., Sampathkumar P., Easton A.M., Luckow V.A., Rownd R.H.;  
 RT "Transcription of the replication control region of the IncFII  
 RT R-plasmid NR1 in vitro and in vivo.";  
 RL J. Mol. Biol. 181:395-410(1985).

RN [6]  
 RP SEQUENCE FROM N.A.  
 RC PLASMID=IncFVI pSU212;  
 RX MEDLINE=91324851; PubMed=1865183;  
 RA Lopez J., Delgado D., Andres I., Ortiz J.M., Rodriguez J.C.;  
 RT "Isolation and evolutionary analysis of a RepFVIB replicon of the  
 RT plasmid pSU212.";  
 RL J. Gen. Microbiol. 137:1093-1099(1991).  
 CC -!- FUNCTION: THIS PROTEIN IS INVOLVED IN THE DETERMINATION OF COPY  
 CC NUMBER IN GENE REPLICATION. IT BINDS TO THE REPA PROMOTER THUS  
 CC INHIBITING THE SYNTHESIS OF THE MRNA FOR THE INITIATOR PROTEIN  
 CC REPA.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; J01762; AAA92255.1; -.  
 DR EMBL; M26840; AAA26065.1; -.  
 DR EMBL; V00318; CAA23608.1; -.  
 DR EMBL; M18273; AAA88332.1; -.  
 DR EMBL; X02302; CAA26165.1; -.  
 DR EMBL; X55895; CAA39380.1; -.  
 DR PIR; A04476; QQECAR.  
 DR PIR; I51821; I51821.  
 KW Plasmid; Plasmid copy control; Transcription regulation; Repressor;  
 KW DNA-binding.  
 FT CONFLICT 3 3 Q -> H (IN REF. 2).  
 FT CONFLICT 77 84 SEAQKRCM -> K (IN REF. 2).  
 FT CONFLICT 81 84 KRCM -> NDACDDGLTFLSVQKISARLLIV (IN  
 FT REF. 1).  
 SQ SEQUENCE 84 AA; 9438 MW; ADC3442D128E2C81 CRC64;

Query Match 4.8%; Score 4; DB 1; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 LSVA 45  
 ||||  
 Db 32 LSVA 35

RESULT 12

DEFI\_TENMO

ID DEFI\_TENMO STANDARD; PRT; 84 AA.  
AC Q27023;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Tenecin 1 precursor.  
OS Tenebrio molitor (Yellow mealworm).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;  
OC Tenebrionidae; Tenebrio.  
OX NCBI\_TaxID=7067;  
RN [1]  
RP SEQUENCE FROM N.A., SEQUENCE OF 49-59, AND DISULFIDE BONDS.  
RC TISSUE=Larval hemolymph;  
RX MEDLINE=95096025; PubMed=7798186;  
RA Moon H.J., Lee S.Y., Kurata S., Natori S., Lee B.L.;  
RT "Purification and molecular cloning of cDNA for an inducible  
RT antibacterial protein from larvae of the coleopteran, Tenebrio  
RT molitor.";  
RL J. Biochem. 116:53-58(1994).  
CC -!- FUNCTION: BACTERICIDAL PROTEIN PRODUCED IN RESPONSE TO INJURY. IT  
CC IS CYTOTOXIC TO GRAM-POSITIVE BACTERIA.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D17670; BAA04552.1; -.  
DR PIR; JX0332; JX0332.  
DR HSSP; P10891; 1ICA.  
DR InterPro; IPR001542; Defensin\_anpod.  
DR InterPro; IPR003614; Knot1.  
DR Pfam; PF01097; Arthro\_defensin; 1.  
DR PRINTS; PR00271; DEFENSIN.  
DR SMART; SM00505; Knot1; 1.  
DR PROSITE; PS00425; ARTHROPOD\_DEFENSINS; 1.  
KW Insect immunity; Antibiotic; Signal.  
FT SIGNAL 1 19 POTENTIAL.  
FT PROPEP 20 41  
FT CHAIN 42 84 TENECIN 1.  
FT DISULFID 44 75  
FT DISULFID 61 81  
FT DISULFID 65 83  
SQ SEQUENCE 84 AA; 9176 MW; 0367C5070468BE60 CRC64;

Query Match 4.8%; Score 4; DB 1; Length 84;  
Best Local Similarity 100.0%; Pred. No: 3.1e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRSG 5  
|||  
Db 69 GRSG 72

RESULT 13

EX7S\_CAUCR

ID EX7S\_CAUCR STANDARD; PRT; 84 AA.  
AC Q9A6M3;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Probable exodeoxyribonuclease VII small subunit (EC 3.1.11.6)  
DE (Exonuclease VII small subunit).  
GN XSEB OR CC2070.  
OS Caulobacter crescentus.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;  
OC Caulobacteraceae; Caulobacter.  
OX NCBI\_TaxID=155892;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 19089 / CB15;  
RX MEDLINE=21173698; PubMed=11259647;  
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
RT "Complete genome sequence of *Caulobacter crescentus*."  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large  
CC acid-insoluble oligonucleotides, which are then degraded further  
CC into small acid-soluble oligonucleotides (By similarity).  
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-  
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.  
CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE XSEB FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE005880; AAK24041.1; -.  
DR PIR; E87505; E87505.  
DR TIGR; CC2070; -.  
DR HAMAP; MF\_00337; -, 1.  
DR InterPro; IPR003761; Exonuc\_VII\_S.

DR Pfam; PF02609; Exonuc\_VII\_S; 1.  
DR ProDom; PD028235; Exonuc\_VII\_S; 1.  
DR TIGRFAMs; TIGR01280; xseB; 1.  
KW Hydrolase; Nuclease; Exonuclease; Complete proteome.  
SQ SEQUENCE 84 AA; 8979 MW; E6519EB31EBACE8A CRC64;

Query Match 4.8%; Score 4; DB 1; Length 84;  
Best Local Similarity 100.0%; Pred. No. 3.1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 EALS 43  
|||  
Db 18 EALS 21

#### RESULT 14

##### EX7S\_HAEIN

ID EX7S\_HAEIN STANDARD; PRT; 84 AA.  
AC P43914;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Exodeoxyribonuclease VII small subunit (EC 3.1.11.6) (Exonuclease VII  
DE small subunit).  
GN XSEB OR HI1437.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Haemophilus.  
OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Rd / KW20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
RT Rd.";  
RL Science 269:496-512(1995).  
CC -!- FUNCTION: BIDIRECTIONALLY DEGRADES SINGLE-STRANDED DNA INTO LARGE  
CC ACID-INSOLUBLE OLIGONUCLEOTIDES, WHICH ARE THEN DEGRADED FURTHER  
CC INTO SMALL ACID-SOLUBLE OLIGONUCLEOTIDES. IT CAN ALSO DEGRADE 3'  
CC OR 5' SS REGIONS EXTENDING FROM THE TERMINI OF DUPLEX DNA  
CC MOLECULES AND DISPLACED SS REGIONS (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-  
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.  
CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE XSEB FAMILY.  
CC -----



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; U32822; AAC23086.1; -.  
DR PIR; A64172; A64172.  
DR TIGR; HI1437; -.  
DR HAMAP; MF\_00337; -, 1.  
DR InterPro; IPR003761; Exonuc\_VII\_S.  
DR Pfam; PF02609; Exonuc\_VII\_S; 1.  
DR ProDom; PD028235; Exonuc\_VII\_S; 1.  
DR TIGRFAMs; TIGR01280; xseB; 1.  
KW Hydrolase; Nuclease; Exonuclease; Complete proteome.  
SQ SEQUENCE 84 AA; 9511 MW; 4A3CF6FF855BA72E CRC64;

Query Match 4.8%; Score 4; DB 1; Length 84;  
Best Local Similarity 100.0%; Pred. No. 3.1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 ASSQ 71  
|||  
Db 6 ASSQ 9

#### RESULT 15

##### GVM1\_HALN1

ID GVM1\_HALN1 STANDARD; PRT; 84 AA.  
AC P24377; Q9HI28;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Gas vesicle protein gvpM 1.  
GN (GVPM11 OR GVPM OR VNG5019G) AND (GVPM12 OR VNG6019G).  
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081), and  
OS Halobacterium sp. (strain NRC-817).  
OG Plasmid pNRC100, Plasmid pNRC200, and Plasmid pHH1.  
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
OC Halobacteriaceae; Halobacterium.  
OX NCBI\_TaxID=64091, 148370;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NRC-1; PLASMID=pNRC100;  
RX MEDLINE=91323716; PubMed=1864501;  
RA Jones J.G., Young D.C., Dassarma S.;  
RT "Structure and organization of the gas vesicle gene cluster on the  
RT Halobacterium halobium plasmid pNRC100.";  
RL Gene 102:117-122(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NRC-1; PLASMID=pNRC100;  
RX MEDLINE=99063795; PubMed=9847077;  
RA Ng W.V., Ciufu S.A., Smith T.M., Bumgarner R.E., Baskin D., Faust J.,  
RA Hall B., Loretz C., Seto J., Slagel J., Hood L., DasSarma S.;

RT "Snapshot of a large dynamic replicon in a halophilic archaeon:  
RT megaplasmid or minichromosome?";  
RL Genome Res. 8:1131-1141(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NRC-1; PLASMID=pNRC200;  
RX MEDLINE=20504483; PubMed=11016950;  
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,  
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,  
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,  
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;  
RT "Genome sequence of Halobacterium species NRC-1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NRC-817; PLASMID=pHH1;  
RX MEDLINE=92065812; PubMed=1956294;  
RA Horne M., Englert C., Wimmer C., Pfeifer F.;  
RT "A DNA region of 9 kbp contains all genes necessary for gas vesicle  
RT synthesis in halophilic archaeobacteria.";  
RL Mol. Microbiol. 5:1159-1174(1991).  
CC -!- FUNCTION: COULD BE IMPORTANT FOR THE SHAPE DETERMINATION OF THE  
CC GAS VESICLE.  
CC -!- SIMILARITY: BELONGS TO THE GAS VESICLE PROTEIN TYPE A FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M58557; AAA98186.1; -.  
DR EMBL; AF016485; AAC82799.1; -.  
DR EMBL; AE005141; AAG20716.1; -.  
DR EMBL; X55648; CAA39180.1; -.  
DR PIR; T08232; T08232.  
DR InterPro; IPR000638; Gas\_vesicle.  
DR Pfam; PF00741; Gas\_vesicle; 1.  
DR ProDom; PD003598; Gas\_vesicle; 1.  
DR PROSITE; PS00234; GAS\_VESICLE\_A\_1; 1.  
DR PROSITE; PS00669; GAS\_VESICLE\_A\_2; 1.  
KW Gas vesicle; Plasmid; Complete proteome.  
SQ SEQUENCE 84 AA; 9248 MW; 4A8F71632E02F6F8 CRC64;

Query Match 4.8%; Score 4; DB 1; Length 84;  
Best Local Similarity 100.0%; Pred. No. 3.1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 SPTA 68  
|||  
Db 76 SPTA 79

Search completed: January 7, 2004, 10:33:40  
Job time : 13 secs